S. Hrg. 103–317

HUMAN GENOME DIVERSITY PROJECT

HEARING

BEFORE THE

COMMITTEE ON GOVERNMENTAL AFFAIRS UNITED STATES SENATE

ONE HUNDRED THIRD CONGRESS

FIRST SESSION

APRIL 26, 1993

Printed for the use of the Committee on Governmental Affairs

U.S. GOVERNMENT PRINTING OFFICE

67–460 cc WASHINGTON : 1993

For sale by the U.S. Government Printing Office
Superintendent of Documents, Congressional Sales Office, Washington, DC 20402
ISBN 0-16-043334-7

S. Hrg. 103–317

HUMAN GENOME DIVERSITY PROJECT

HEARING

BEFORE THE

COMMITTEE ON GOVERNMENTAL AFFAIRS UNITED STATES SENATE

ONE HUNDRED THIRD CONGRESS

FIRST SESSION

APRIL 26, 1993

Printed for the use of the Committee on Governmental Affairs

U.S. GOVERNMENT PRINTING OFFICE

67–460 cc WASHINGTON : 1993

For sale by the U.S. Government Printing Office
Superintendent of Documents, Congressional Sales Office, Washington, DC 20402
ISBN 0-16-043334-7

COMMITTEE ON GOVERNMENTAL AFFAIRS

JOHN GLENN, Ohio, *Chairman*

SAM NUNN, Georgia
CARL LEVIN, Michigan
JIM SASSER, Tennessee
DAVID PRYOR, Arkansas
JOSEPH I. LIEBERMAN, Connecticut
DANIEL K. AKAKA, Hawaii
BYRON L. DORGAN, North Dakota

WILLIAM V. ROTH, JR., Delaware
TED STEVENS, Alaska
WILLIAM S. COHEN, Maine
THAD COCHRAN, Mississippi
JOHN McCAIN, Arizona

Leonard Weiss, *Staff Director*
Shane Merz, *Professional Staff*
Franklin G. Polk, *Minority Staff Director and Chief Counsel*
Michal Sue Prosser, *Chief Clerk*

CONTENTS

HUMAN GENOME DIVERSITY PROJECT

MONDAY, APRIL 26, 1993

U.S. SENATE,
COMMITTEE ON GOVERNMENTAL AFFAIRS,
Washington, DC.

The Committee met, pursuant to notice, at 2:11 p.m., in room SD–342, Dirksen Senate Office Building, Hon. Daniel K. Akaka, presiding.

Present: Senator Akaka.

OPENING STATEMENT OF SENATOR AKAKA

Senator AKAKA [presiding]. Aloha and good afternoon. This hearing of the Governmental Affairs Committee will come to order.

Today, the Committee will hear testimony on a very ambitious project known as the Human Genome Initiative—and a proposed supplemental project known as the Human Genome Diversity Project.

The Human Genome project is a worldwide research initiative to analyze the structure of human DNA and determine the location of the tens of thousands of human genes. The information generated by the Human Genome Project will be of immense benefit to medical science.

The Federal commitment to the project is considerable. The project is estimated to cost more than $3 billion. Given the planned duration and the resources necessary for such an undertaking, a project of this magnitude must be dynamic, and research priorities must constantly be reviewed. When promising research opportunities develop, there must be a mechanism for consideration of supplemental programs.

One such supplemental program is the Human Genome Diversity Project. Throughout the world, many ethnic and aboriginal groups are in danger of losing their genetic identity. There is much that we could learn from these diverse cultural groups. The answers to many anthropological mysteries are hidden in the genes of these ancestral populations. Unfortunately, time is running out. The day may soon come when the richness of our genetic diversity is lost forever. We may lose the opportunity to learn about ourselves from the study of diverse populations.

The issue to be addressed in this hearing is the potential benefit of the Human Genome Diversity Project to medical science. I am also sympathetic to cultural and social science investigations that increase our understanding of native peoples.

I welcome the thoughtful comments of our witnesses and hope that this hearing will highlight many of the scientifically meaning-

ful accomplishments being made in the study of the human genome.

The hearing record will remain open until the close of business May 7, 1993 to provide an opportunity for comments from individuals and organizations that are not present here today.

I ask that each witness limit their testimony to 5 minutes to provide greater time for discussion. Your written statements will be entered into the record in their entirety.

PREPARED STATEMENT OF SENATOR AKAKA

Aloha and good afternoon. This hearing of the Governmental Affairs Committee will come to order.

Today, the committee will hear testimony on an ambitious project known as the Human Genome Initiative and a proposed supplemental project known as the Human Genome *Diversity* Project. The Human Genome Project is a worldwide research initiative to analyze the structure of human DNA and determine the location of the tens of thousands of human genes. The information generated by the Human Genome Project will be of immense benefit to medical science.

The Federal commitment to the project is considerable, the project is estimated to cost more than $3 billion. Given the planned duration and the resources necessary for such an undertaking, a project of this magnitude must be dynamic and it is important to constantly review research priorities. When promising research opportunities develop, there must be a mechanism for consideration of supplemental programs.

One such supplemental program is the Human Genome Diversity Project. Throughout the world, many ethnic and aboriginal groups are in danger of losing their genetic identity. There is much that we could learn from these diverse cultural groups. The answers to many anthropological mysteries are hidden in the genes of these ancestral populations. Unfortunately, time is running out. The day may soon come when the richness of our genetic diversity is lost forever. We may lose the opportunity to learn about ourselves from the study of diverse populations.

An important issue to be addressed in this hearing is the potential benefit of the Human Genome Diversity Project to the medical sciences. I am also sympathetic to cultural and social science investigations that increase our understanding of native peoples.

I welcome the thoughtful comments of our witnesses and hope that this hearing will highlight many of the scientifically meaningful accomplishments being made in the study of the human genome.

The hearing record will remain open until the close of business May 7, 1993 to provide an opportunity for comments from individuals and organizations that are not present today.

I ask that each witness limit their testimony to 5 minutes to provide greater time for discussion. Your written statements will be entered into the record in their entirety.

Senator AKAKA. I'd like to call our witnesses forward. There will be two panels. The first panel will consist of the following—and as I call you, will you please come up to the witness table—Dr. Cora Marrett, Assistant Director for Social, Behavioral and Economic Research, National Science Foundation; Dr. Francis Collins, National Center for Human Genome Research, National Institutes of Health; Dr. David J. Galas, Associate Director for Health and Environmental Research, U.S. Department of Energy; and Dr. Robyn Nishimi, Biological and Behavioral Sciences Program, Office of Technology Assessment.

I would also like to welcome Dr. Ruth Kirschstein, Director of the National Institute of General Medical Science, who is in the audience. Aloha and welcome. Dr. Kirschstein has kindly agreed to serve as a resource witness.

At this time, in the order that I called you, please proceed.

Dr. Marrett?

TESTIMONY OF DR. CORA MARRETT, ASSISTANT DIRECTOR FOR SOCIAL, BEHAVIORAL AND ECONOMIC RESEARCH, NATIONAL SCIENCE FOUNDATION

Dr. MARRETT. Thank you very much, Mr. Chairman.

I appreciate very much the opportunity to appear before you today with my colleagues here at the table and in the audience to discuss the Human Genome Diversity Project.

Recent remarkable advances in molecular genetics have enabled the launching of the Human Genome Project, whose mission is the decoding of the total package of genetic material, or DNA, in a human. This massive effort is well underway and is almost 5 years old. It is funded by the National Institutes of Health and the Department of Energy and currently has an annual budget of approximately $180 million. It is still many years away from completion.

Once the Genome Project produces an entire DNA sequence, we will still not have extensive information on the hereditary diversity that exists among people—the very large differences that exist among people within ethnic groups or populations, and the smaller differences that characterize differences among ethnic groups or populations.

The Human Genome Diversity Proposal is a developing effort of a group of physical and cultural anthropologists, archaeologists, geneticists, epidemiologists, and linguists to collect, organize, and analyze sufficient DNA from different people and populations to answer two major issues.

The first issue is understanding human genetic change. By laying out the specific DNA differences among individuals worldwide, it will be possible to reconstruct the relationships and origins of the different human populations—the ancient origins of different Native American groups, for example, or Polynesians, of different European and African groups, and so on.

The second issue is the hereditary basis for differences in human susceptibility to diseases. We are learning weekly the genetic determination of many comparatively rare diseases—Huntington's Chorea, cystic fibrosis, and so on. The list is now thousands long. For a very few of these, we know the exact location and DNA sequence of the gene involved.

In addition, we know that the more common diseases, afflicting millions of people, often involve both genetic predispositions and environmental effects. Well-established cases include breast cancer, diabetes, bowel cancer, and at least certain cardiovascular conditions. These problems and others as well vary in their risk from one family to the next and from one ethnic group to the next. Heredity clearly plays a role. Learning more about the gene or genes that place people at higher risk for these will inevitably help in their treatment and prevention.

I want to stress that the Human Genome Diversity Project is at this point still only a proposal. The Human Genome Organization and international scientific organizations established a committee in 1991 to develop a Human Genome Diversity Project. As a part of that effort, the committee in the United States submitted a proposal to the physical anthropology program at the National Science Foundation to support a series of workshops that would develop a plan for amassing genetic, linguistic, medical and environmental

information on a diverse group of human populations from around the world. More on the workshops will be filled in by others in this panel and the one to follow.

The physical anthropology program at the National Science Foundation coordinated modest levels of funding for these workshops with other programs from the National Science Foundation, with the Center for Human Genome Research, the National Institute of General Medical Sciences, and the Department of Energy.

Three of the four planned workshops have now been held. The first focused on what an adequate sample of people would be to characterize a population. The second focused on what particular populations were best candidates for inclusion in the project. Over 400 populations were identified. The third focused on ethical issues, biotechnical developments and funding possibilities. The fourth will involve international possibilities for collaboration and funding.

Since the project clearly must be an international one, there has been an attempt from the outset to enlist cooperation and parallel funding from other countries. European members of the Human Genome Organization Committee have worked to establish funding to support the Human Genome Diversity Project and are working to allow funding for European components of the project. They have secured some funding to begin work, and we can say more about the international funding, again, at a later point.

The structure of the project that is evolving will have at its center an international Human Genome Diversity Scientific Committee. A second committee of importance will be the Ethics Committee, a committee to monitor, assess and anticipate the complex ethical issues that inevitably will develop.

The project then can be thought of as divided into three parts. The first is the collection phase, involving the identification of appropriate populations or groups to be covered in the survey; the collection of required family, medical, demographic, cultural and ecological information, and the taking of blood samples from cooperating individuals. At this point, at least 400 groups worldwide are likely to be included.

The second part involves the transformation and storage of the blood samples themselves in central laboratories. Along with the other data collected on the populations, stored in computer files, these will make a permanent and inexhaustible source of information for future scientific and health research.

The third part is the analysis phase, which will extend for some time into the future.

As we see it, this is an idea whose time has come. You mentioned earlier the rapidity with which diversity is disappearing, being eroded and obscured by a variety of forces as there is the penetration of heretofore isolated ecologies and environments, and changing diets and lifestyles accompanying the alterations that take place.

We need a systematic collection of information, genetic and otherwise, on the diversity of human populations before they become even more diluted by the growing global mobility of peoples. This will be an invaluable resource for understanding our origins and

our hereditary accommodations and susceptibilities to the diseases that afflict us.

Thank you, Mr. Chairman. I'd be pleased to answer questions at the time you'd like to raise them.

Senator AKAKA. Thank you very much, Dr. Marrett.

PREPARED STATEMENT OF DR. CORA MARRETT

Mr. Chairman and Members of the Committee, I appreciate very much the opportunity to appear before you today with my colleagues from the National Institutes of Health, the Energy Department, and the Office of Technology Assessment to discuss the Human Genome Diversity Project.

RELATION TO THE HUMAN GENOME PROJECT

Recent remarkable advances in molecular genetics have enabled the launching of the Human Genome Project, whose mission was the decoding of the total package of genetic material, or DNA, in a human. This massive effort is well underway and is almost five years old. It is funded by the National Institutes of Health and the Department of Energy, and currently has an annual budget of approximately $180 million. It is still many years away from completion.

Once the Genome Project produces an entire DNA sequence, we will still not have extensive information on the hereditary *diversity* that exists among people—the very large differences that exist among people within ethnic groups or populations, and the smaller differences that characterize differences among ethnic groups or populations.

The Human Genome Diversity Proposal is a developing effort of a group of physical and cultural anthropologists, archaeologists, geneticists, epidemiologists, and linguists to collect, organize, and analyze sufficient DNA from different people and populations around the world to answer two major issues.

OBJECTIVES OF THE HUMAN GENOME DIVERSITY PROJECT

The first issue is understanding human genetic change. By laying out the specific DNA differences among individuals world-wide, it will be possible to reconstruct the relationships and origins of the different human populations—the ancient origins of different Native American groups, for example, of Polynesians, of different European and African groups, and so forth.

The second issue is the hereditary basis for differences in human susceptibility to diseases. We are learning weekly the genetic determination of a many comparatively rare diseases—Huntington's Chorea, cystic fibrosis, and so on. The list is now thousands long. For a very few of these, we know the exact location and DNA sequence of the gene involved.

In addition, we know that the more common diseases, afflicting millions of people, often involve both genetic predispositions and environmental effects—well-established cases include breast cancer, diabetes, bowel cancer, and at least certain cardiovascular conditions. These problems, and others as well, vary in their risk from one family to the next and one ethnic group to the next. Heredity clearly plays a role. Learning more about the gene or genes that place people at higher risk for these will inevitably help in their treatment and prevention.

CURRENT STATUS OF THE HUMAN GENOME DIVERSITY PROPOSAL

I want to stress that The Human Genome Diversity Project is, at this point, still only a proposal. The Human Genome Organization (HUGO), and international scientific organizations, established a committee in 1991 to develop a Human Genome Diversity Project. As part of that effort, the committee members in the United States, including Dr. Cavalli-Sforza and Dr. King, submitted a proposal to the Physical Anthropology Program at NSF to support a series of workshops that would develop a plan for amassing genetic, linguistic, medical, and environmental information on a diverse group of human populations from around the world. A NSF Physical Anthropology Program coordinated modest levels of funding for these workshops with selected other NSF programs, with the Center for Human Genome Research, the National Institute of General Medical Sciences, and the Department of Energy.

Three of four planned workshops have now been held. The first focused on what an adequate sample of people would be to characterize a population. The second focussed on what particular populations were best candidates for inclusion in the project. Over 400 populations were identified. The third focussed on issues,

biotechnical developments, and funding possibilities. The fourth will involve international possibilities for collaboration and funding.

Since the project clearly must be an international one, there has been an attempt from the outset to enlist cooperation and parallel funding from other countries. European members of the HUGO committee have worked to establish funding to support the HGD project, and are working to allow funding for European components of the HGD project They have already secured some funding to begin work. Technical and scientific representatives of the E.C. and Japan have also been invited to the workshops, and individual scientists from a number of countries around the world have participated in the workshops. Certain private agencies likely to be interested in the project have also been approached by the organizers.

The structure of the project that is evolving will have, at its center, an International HGD Scientific Committee. A second committee of importance will be the HGD Ethics Committee, to monitor, assess and anticipate the complex ethical issues that will inevitably develop.

The project itself can be thought of as being divided into three parts.

The first part will be the collection phase. This involves the identification of appropriate populations or groups to be covered in the survey; the collection of required family, medical, demographic, cultural and ecological information (on each group if it has not already been made); and the taking of blood samples from cooperating individuals. At this point, at least 400 groups world-wide are likely to be included.

The second part involves the transformation and storage of the blood samples themselves in central laboratories. Along with the other data collected on the populations, stored in computer files, these will make a permanent and inexhaustible source of information for future scientific and health research.

The third part is the analysis phase, which will extend for some time into the future.

SUMMARY

This is an idea whose time has come. The great hereditary diversity of humanity is rapidly being eroded or obscured by the same forces that are transforming so much of our world today—increased ease of movement and migration, the penetration of heretofore isolated ecologies and environments by the world economy and its representatives, and changing diets and life styles accompanying these alterations. We need a systematic collection of information, genetic and otherwise, on the diversity of human populations before they become even more diluted by the growing global mobility of peoples. This will be an invaluable resource for understanding our origins and our hereditary accommodations and susceptibilities to the diseases that afflict us.

Senator AKAKA. Dr. Francis Collins, we'd be pleased to have your testimony.

TESTIMONY OF DR. FRANCIS COLLINS, DIRECTOR, NATIONAL CENTER FOR HUMAN GENOME RESEARCH, NATIONAL INSTITUTES OF HEALTH

Dr. COLLINS. Thank you, Mr. Chairman.

I am happy to address this issue in my first appearance before a congressional committee since becoming Director of the National Center for Human Genome Research on April 4th.

I am a physician and a geneticist. I have been involved, prior to taking on this position, in the search for genes that cause human disease, particularly cystic fibrosis, neurofibromatosis, Huntington's disease, and breast cancer. I am fascinated by the opportunities that genetic research presents for a better understanding of ourselves, for improving our health, and for reducing health care costs.

I want to distinguish between the Human Genome Project (which is now underway, supported by a combination of sources including the NIH and the Department of Energy) and the Diversity Project, which is the topic of today's discussion. As Dr. Marrett has said, the Diversity Project is a proposal seeking a direction.

The Human Genome Project, I believe, is the most important, exciting, and potentially beneficial scientific effort that humankind has ever undertaken. I know that sounds like a very strong statement, but I think it is quite defensible.

The map and the sequence of all 100,000 human genes, which we expect to have by the year 2005, will tell us the basis of the 4,000 or so genetic diseases which afflict us. They will also tell us about the genetic basis of more common disorders like diabetes and hypertension. Almost no one has a family history that is completely without examples of familial illnesses.

This genetic information will lead us to much better diagnostics and to treatments that we haven't dreamed of before. For cystic fibrosis, for example, the finding of the gene 3 years ago has led to the initiation of gene therapy trials within the past 2 weeks. This is a very telescoped time period, and such examples will be seen again as we discover other genes for genetic disease.

Huntington's disease has already been mentioned. After a long search, that gene has emerged within the last month. The rate of acceleration of genetic disease gene identification is really quite dramatic.

One of the beauties of the Human Genome Project is that it is already producing these kinds of results early on in its progress; you don't have to wait until the end to enjoy the benefits.

A gene for breast cancer is being avidly sought in a collaboration between my research group and that of Dr. Mary Claire King, one of the later witnesses, as well as a number of other groups around the world. That gene on Chromosome 17, is carried by one in 200 women and if you have it, places you at a very high risk of breast cancer. That gene will probably be found within the next year. The identification of that gene and the ability to identify women who are at very high risk and get them into intense surveillance programs promises to do much for this terrible disorder that strikes so many. In itself this finding may be able to actually pay back the costs of the Genome Project which are, I would argue, rather modest for the benefits that the project will produce.

I think it is useful to make a quick comment on the history of the Human Genome Project as we now look at the Diversity Project and try to plan the best way to move forward. Let me say at the outset that I support the Diversity Project. I think it is useful, however, to consider the best way to get a very complicated scientific endeavor of this sort, with many ethical dimensions, underway.

For the Human Genome Project, discussions about carrying out this activity began in about 1985. In a series of scientific meetings, ideas about how the project might be organized were put forward. The project's goals modulated considerably during those 2 or 3 years. In my view, a very important part of that process was an external review of the Genome Project by the National Research Council and another one by the OTA, represented here today. That provided an opportunity for those who were not themselves in the middle of the effort to stand back and look at the goals, to think carefully about the budget, and to make suggestions as to the best organizational structure to get the job done.

Those reports were carefully reviewed, and when the Genome Project's organization was set up, it was done with detailed goals

and timetables. We have adhered to these very tightly since that time. The Genome Project is thus a very focused effort. It is focused on developing genetic and physical maps, and the sequences of the human genome, as well as that of model organisms that tell us more information about the human genome.

Partly as a result of the advice received in those early stages, a strong component of the Human Genome Project is the Ethical, Legal and Social Implications (ELSI) Branch. This Branch has been devoting much attention to the consequences of having increased ability to obtain genetic information on individuals in the population and particularly how to avoid the chance that people will be discriminated against on the basis of genetic information that is determined about them. Five percent of the genome budget is now devoted to this ELSI program, and most observers think it is a very crucial part of the whole endeavor.

The Genome Project is going extremely well. The goals that we set for ourselves in genetic and physical mapping are actually being exceeded despite the fact that the budget for this effort has never achieved the full ramp up that was suggested by the original review panels.

So we are running lean, but determined to succeed, and we really have a lot on our plate to get this job done.

I support the Diversity Project as a fascinating scientific endeavor, with great consequences for our understanding of the relatedness and origins of human populations. I believe that we will learn things about medical disorders as well, although those consequences are a little less easy to be certain of at the present time and will require further study. The present plan calls for collecting blood samples from about 25 members of each of 400 populations. Unless a particular disease gene is rather frequent in a particular population, there is a good chance it would not be seen in such a small number of blood samples. But obviously this is a trade-off between cost and benefit.

I am concerned, however, about the social implications of this project, and I'm sure that will be discussed this afternoon. Will the Third World, in which many of these populations currently reside, feel exploited by the project's coming in and obtaining samples from such individuals? I think it is crucial to obtain consultation at an early stage from representatives of that community to anticipate the kinds of reactions which will come about.

How will we handle informed consent in situations where the usual U.S. Government rules may be difficult to put into place as regards the written aspects of informed consent? Perhaps most importantly, how will we guard against misuse of this information? By determining information about ethnic groups that characterizes them and differentiates them from each other, there will be the possibility of demagogues taking that information and using it to discriminate against ethnic groups. I believe that we can protect against that outcome if we prepare carefully for it, but I think a great deal of attention needs to be paid to that issue before we get too far along.

For those reasons, I believe that the experience gained through setting up the Human Genome Project applies to the Diversity Project, and that it would be very appropriate at this juncture to

have a review, perhaps by a body like the NRC, of this project, to make specific recommendations about goals, budgets, and how to handle the very thorny ethical issues.

I appreciate the chance to offer these comments, and I would be happy to answer questions.

Senator AKAKA. Thank you very much, Dr. Collins.

PREPARED STATEMENT OF DR. FRANCIS COLLINS

Mr. Chairman, Senator Akaka, and Members of the Committee, it is my pleasure to appear before you today as the new Director of the National Center for Human Genome Research (NCHGR) of the National Institutes of Health (NIH) and to have the opportunity to talk to you about the Human Genome Project. I am enormously pleased to be at the helm of what I consider to be the single most important scientific endeavor we have ever embarked upon. By the end of the 15-year project, we hope to have produced detailed maps of all the human chromosomes and determined the sequence of the 3 billion pairs of nucleotide bases that make up human DNA. This information will be stored in databases that will allow researchers to have access to any region of the human genome right at their fingertips. The multitude of benefits we are witnessing already from the Human Genome Project is only the beginning of what I believe will be a revolution in molecular medicine and human biology.

Goals of the Human Genome Project

The Human Genome Project is an international research effort that has the goal of analyzing the structure of human DNA and determining the location of the estimated 100,000 genes located on 23 pairs of human chromosomes. In the United States, the Human Genome Project is managed principally by two government agencies. the NIH and the Department of Energy (DOE), who together have set specific scientific goals in an initial 5-year plan to manage this historic effort. You will also be hearing testimony today from Dr. David Galas of the DOE.

The primary mission of the Human Genome Project is to develop research tools—chromosome maps, DNA sequence information, laboratory technology, and computer databases—that will allow researchers to find and analyze genes faster, more easily, and more cheaply. These tools will have tremendous benefits for biomedical research and make important contributions to a variety of research projects. Of primary interest are the extraordinary medical benefits that will result from our ability to understand the genetic factors of health and disease. Gene discovery gives researchers the opportunity to study the function of the gene and its role in cell biology. This knowledge will revolutionize our strategies to diagnose, treat, and even prevent many diseases.

Medical Benefits

The pace of disease gene discovery has increased substantially because of the research tools developed by the Human Genome Project. We have seen the discovery of genes responsible for genetic diseases such as cystic fibrosis, neurofibromatosis type I and II, fragile X, and most recently, you may have heard about the discovery of the gene responsible for Huntington's disease. In collaboration with another of today's witnesses, Dr. Mary-Claire King, my colleagues and I are also zeroing in on a gene that causes breast cancer—I expect that the gene will be located within the next year. It is a gene that 1 out of 200 women inherit; these women have an 85 percent chance of getting breast cancer and have an increased risk of ovarian cancer as well. We hope soon to be able to offer diagnostic testing for women at risk, and eventually develop ways to treat and prevent this type of early-onset breast cancer.

The isolation of the cystic fibrosis gene provides a good example of the progress that can be made in understanding a disease once the gene is isolated. I was part of a team of researchers who isolated the gene in 1989, and I am excited to report that we have already begun to design new, highly specific drugs to treat cystic fibrosis and are in the process of developing gene therapy techniques. The first human gene therapy experiment targeting cystic fibrosis began this month on the NIH campus by researchers at the National Heart, Lung and Blood Institute. The ability to begin gene therapy trials less than 4 years after the gene was discovered is a dramatic example of the power of the new tools of human molecular genetics. NIH has recently established a Division of Intramural Research within the NCHGR which will focus on technologies for finding disease genes and developing subsequent DNA diagnostics and gene therapies.

The Human Genome Project is cutting edge basic science, and is providing the research tools to accelerate our understanding of the biological and molecular processes that permit human life to develop and function. This will profoundly affect our ability to understand the molecular basis of disease and will greatly improve our ability to diagnose, treat, and prevent many common diseases resulting from malfunctions or mutations in our genes. Such diseases represent a major fraction of the chronic conditions that account for most of the health care costs today.

The Human Genome Diversity Project

The Human Genome Diversity Project, which we are here to discuss today, is one of the many research projects that will be greatly benefitted by the genome analysis tools being developed by the Human Genome Project. The objectives of the Human Genome Diversity Project are to collect, analyze, and preserve genetic samples from a host of vanishing human populations. This project has the potential of giving us new knowledge about human origins, evolutionary history, and genetic diversity; it may also eventually lead to a better understanding of the frequency and susceptibility to disease among diverse populations. It is timely to discuss the undertaking of the Human Genome Diversity Project; I am mindful that researchers wish to collect the samples now before we lose the opportunity due to the further breakdown of geographical barriers, war, famine, or disease.

The Human Genome Diversity Project is beyond the mission of the federally funded Human Genome Project. The NIH and DOE are facing challenges to accomplish the goals we set out to achieve within the next 15 years. We have a long road ahead to complete the genetic and physical maps of all the human chromosomes, and it is essential that there be further improvements in sequencing technology if we wish to sequence the entire human genome quickly and efficiently. We have been able to make significant progress because we have clearly defined our mission. At this time, the Human Genome Project's major contribution to the Human Genome Diversity Project would be for us to continue to develop the research tools that will allow genetic diversity research, as well as other scientific research disciplines, to proceed in a cost-effective way. There may be aspects of technology development, for example high throughput genotyping of many DNA samples, which will benefit both the Human Genome Project and genetic diversity studies, and we look forward to exploring those complementary areas.

I believe the Human Genome Diversity Project is a valuable international endeavor, and that the NIH should cooperate in this project. We are pleased that the NCHGR was able to participate in and contribute funding to the National Science Foundation's grant to the Human Genome Organization (HUGO) Committee for Human Genetic Diversity for a series of workshops on the Human Genome Diversity Project. The National Institute of General Medical Sciences at NIH and the DOE were also contributors. The most recent workshop took place on February 16–18, 1993, on the NIH campus in Bethesda, Maryland, and it included discussions on the ethical issues surrounding genetic diversity research.

Ethical, Legal, and Social Issues

The information generated by the Human Genome Diversity Project will further our knowledge about our human origins and evolution, but it also will raise some challenging ethical, legal, and social issues that must be identified and addressed before the project begins. The February workshop raised several of these important issues including: (1) the ethical issues raised in doing biomedical research in developing countries and insuring the protection of human subjects (for example, it is difficult to ascertain what "informed consent" means in other cultures to know if an individual is voluntarily participating in the sample collection process); (2) the legal issues raised by the possible commercial value of the project's samples or results; and (3) the social-political issues surrounding the possible misuse or misinterpretation of the information generated. Research concerning human genetic differences always merits careful attention to avoid notions of superiority or inferiority among diverse populations.

The Human Genome Project has also faced challenging issues related to the use of human genetic information. At the NCHGR, the Ethical, Legal, and Social Implications Branch was created to define these issues and to develop initial policy options to address them. Five percent of our budget is devoted to the activities of this branch. Our experience has shown that the need is great to examine the ethical, legal, and social issues—alongside with the scientific research—in order to minimize any adverse social consequences resulting from the generation of genetic information.

I would recommend that the Human Genome Diversity Project, as it is organized and funded, develop mechanisms to address the ethical and legal challenges the project will encounter. I would further suggest that an advisory group be established

to monitor the research and the use of the information it generates. My staff and I would be happy to provide consultation and advice based on our own experience in this endeavor.

I would be pleased to answer any questions that you may have.

Senator AKAKA. Dr. Galas?

TESTIMONY OF DR. DAVID J. GALAS, ASSOCIATE DIRECTOR FOR HEALTH AND ENVIRONMENTAL RESEARCH, U.S. DEPARTMENT OF ENERGY

Dr. GALAS. Thank you, Mr. Chairman.

I am honored and pleased to be here with you today to discuss the Human Genome Project and the issue on the table. My colleagues have been rather eloquent, particularly Francis in his description of the Genome Project, and I can probably be a bit briefer as a result.

The scientific questions that are concerned with the genetic diversity of humankind are indeed extremely important questions—there is no question about that—and all the more fascinating because they really deal with who we are and where we came from. So as to the issues at hand, I don't think anyone would disagree with the statement that these are extremely important and very interesting questions that lie before us in the unknown area of scientific investigation.

Now, the important issue for us here is the relationship of this proposed effort to explore genome diversity to the Human Genome Project. I am delighted that I can report to you in much the same tone that Francis began that the Human Genome Project is now going extremely well. Progress is outstanding, and the cooperation between the National Institutes of Health and the Department of Energy is increasing, and I have every confidence that in spite of the fact that the budgetary projections which lead to the number of $3 billion have not been met, we are still doing extremely well both in the areas of technology development as well as in the area of actually beginning to develop maps and doing things that perhaps a couple of years ago, when we began planning this, we would not have thought we would be able to do.

But we have a long way to go; there is no question about that. We are well-launched on this journey, but it is only beginning.

So to the relationship between the Human Genome Diversity Proposal and the Human Genome Project, it is clear that there are many potential relationships, but I think the key things are that the Human Genome Project will be able to provide technology and scientific information and be a scientific resource for much of future biology and medicine. In fact, it is just in those terms that the project was framed, and indeed, all of the thinking that went into that has been borne out in my view that the potential for the scientific basis of biology in the 21st century and in the practice of medicine, the development of medical technology, is an extraordinary benefit of the Human Genome Project as it stands. The development of the Human Genome Project as a scientific resources, I think, is a very important concept here.

So the medical implications of the Human Genome Project, doing all of medical genetics, medical technology development, must be viewed as being very important spinoffs or important supplemental

projects that already exist to the Human Genome Project because it is indeed the foundation of much of future biology and medicine.

The other comment I would like to make is that I, too, am very concerned about the potential ethical, legal and social issues that are raised by the proposals for the Human Genome Diversity Project, and I don't think anyone would slight the importance of those. Those are things that have to be looked at extremely carefully. Misuse of genetic information is something that we have not dealt with well yet in this country from a legal and policy point of view, and the project, the NIH/DOE Human Genome Project, has been looking very carefully at these issues, and I think at least some of the benefits of those studies and the concerns that have been raised as a result should go into a very careful planning process to doing something like a Human Genome Diversity Project.

The final comment I would have is that I think this project, much like Francis said, must stand on its merits, scientifically, insofar as its goals are concerned, whether they be anthropological, linguistic, medical or whatever; it must be reviewed extremely carefully, and if that is done, and if it is planned in such a way that it maximizes the benefit to mankind both from the point of view of scientific information and from the potential medical benefits—and this is reviewed very carefully on a sound scientific basis—I am sure that the Human Genome Project can provide a great deal of the technological knowledge and scientific information that could make it a success.

Thank you very much. I'd be happy to answer questions.

Senator AKAKA. Thank you, Dr. Galas.

PREPARED STATEMENT OF DR. DAVID J. GALAS

I am very pleased to have the opportunity to discuss the historic Human Genome Project with you today. In many ways, the Human Genome Project epitomizes the promise of the biological sciences for the future. The impact on our lives of the vast amount of new knowledge of the living world is just beginning to be realized—it will inevitably transform the biological sciences, medicine, as well as agriculture, food manufacture, chemical technology, and many other areas in ways that are difficult to predict. The project is even now producing a strong shift in the biomedical sciences, from descriptive phenomenology to fundamental understanding, and changes in the practice of medicine will not be far behind. One of the most significant of these latter changes will be a shift in emphasis from curative to preventative medicine—from therapy to prevention. This shift will soon begin to occur because the research that focuses on the fundamental genetic determinants of the functioning of the human body will enable more and more detailed understanding of the nature and causes of human diseases. This will, in turn, allow effective prevention, early intervention and more benign and less costly treatment. This knowledge will include an understanding of the propensities of different individuals for different diseases and susceptibilities to health-endangering practices and environmental influences.

The Department of Energy (DOE) has sponsored the Human Genome Project (HGP) from its inception. At that time, it was recognized that the essential technology which would enable us to undertake this historic project was within our reach, and that immense potential benefits to biology and medicine could be realized. The DOE and the National Institutes of Health (NIH) jointly support this national research effort, carefully coordinating our respective programs.

The HGP is a project that also holds within its promise some significant issues and questions concerning the nature and uses of human genetic information. While the promise of human genome research is enormous, some of the ethical, legal, and social issues associated with this new knowledge are serious and difficult. The HGP has supported an unprecedented effort—more than 3 percent of its budget is devoted to the effort—to study these issues and examine the ethical, legal and social consequences of the coming deluge of genetic information and the new, associated tech-

nologies. These issues include the proper legal restraints on discrimination based on genetic information and the rights to privacy of genetic information and their consequences. It has been well recognized that ethical, legal, and social impacts will accompany greater knowledge about the human genome and that some attempts need to be made to anticipate, and to try to ameliorate, these impacts. Another issue has to do with the proper distribution for the public good of intellectual property rights for genetic material and information—patents on genes.

We consider this historic project to be of immense potential value to the American people and to the world. Thus, we strongly support the expeditious transfer of the benefits of this research and of the associated technologies to the public, which of necessity involves the patent system. This issue has been discussed elsewhere, and I will not comment further here on the gene patenting issue.

Goals and Progress of the Human Genome Project

The general approach to elucidating the genetic contents of the genome is to physically "map" the chromosomes and then sequence large parts of it. Mapping one of the 24 distinct human chromosomes means producing a linear series of DNA fragments, containing genes, that extend collectively from one end of that chromosome to the other. This would then enable the location, isolation and characterization of all of the individual genes and functional sites in that chromosome. The over-arching goal of the HGP is to create a knowledge base of unprecedented detail and complexity as a resource for scientific investigations that will enable subsequent research to be immensely more effective and efficient. The process of mapping chromosomes is well underway and, it is fair to say, is making spectacular progress. Never before has so much genetic information been gained, and so many genes located, identified or characterized so rapidly. Never before has the technical means to gain information been more promising. Never before has so much biological knowledge been generated as during the past few years of genome research. No matter how strongly I emphasize it, you will undoubtedly be surprised in the next few years by the sheer rate at which new information is produced. Let me cite a few recent examples to illustrate this point. Hardly a week goes by now that a discovery in human genetics, relevant to some heritable disease, does not appear in the popular press. Recently, genes involved in muscular dystrophy, fragile-X syndrome, Huntington's disease, "Lou Gehrig's disease" and many others have been discovered because of advances in human genetic knowledge. These discoveries are the harbingers of much more new knowledge that will be gained because of the HGP. The resource of knowledge of genomic maps, for example, now allows disease-related genes to be found much more efficiently. This prospect of new knowledge offers great promise and opportunity to many different branches and disciplines of the biological and medical sciences.

The original and explicit intent for the Human Genome Project (as arrived at by the DOE Office of Energy Research and the NIH National Center for Human Genome Research, and defined in the 1988 DOE–NIH Memorandum of Understanding and the interagency 5-year plan which commenced on October 1, 1990), is clear. Broadly speaking, the goals are twofold and include the mapping and sequencing of the entire human genome, along with the development of advanced technologies and instrumentation to achieve these ends. They also include the development of the informatics capabilities to manage, access, and analyze the resulting data. It is the aim of the U.S. Human Genome Project to accelerate future biological science by building the tools, material resources, and infrastructure so that other branches of science, to be determined by the aggressiveness and imagination of their practitioners, can make the greatest use of these new methods and tools. In other words, the HGP is directed at building a scientific resource of unprecedented complexity and power. Many scientific research enterprises will make use of this resource: medical genetics, for the diagnosis and treatment of disease; and fundamental biology, for the detailed understanding of biological function, to give two examples. Surely, population genetics and anthropology will also be among the early exploiters of the new genetics made possible by the HGP.

Human Genetic Diversity

Let me turn now to the specific emphasis of the present hearing—the question of the nature of human genetic diversity and the research intended to elucidate it. The fundamental scientific questions concerning the nature and full extent of human genetic diversity are very important. The origins and histories of human populations, and the extent to which the genome can and does vary among the present human population, are momentous and fascinating scientific issues. The area of human molecular population genetics has recently attracted much attention, and the opportunities facilitated by new methods, technologies and approaches made possible by advances in molecular biology and genetics are indeed exciting

prospects. It is now possible to distinguish the relatively distant genetic origins of individuals, in addition to being able to distinguish individuals from one other. The former is based on differences in the human genome that persist throughout large groups of people with common heritage. The latter is sometimes referred to as DNA forensics, and is based on relatively recent genetic changes that result in differences among individuals. It is important to recognize, however, the natural boundary between the building of a powerful scientific resource, which is the HGP, and a broad field of scientific investigation that encompasses physical anthropology, human evolution and population genetics.

For several reasons, it would be a serious mistake to ignore this distinction and attempt to bring this sort of research initiative under the umbrella of the HGP. The importance and breadth of the questions of human genetic diversity are such that this research should be constituted as an effort that is clearly distinct from the HGP, both in organization and funding. The most direct reason is that the goals of the research are clearly distinct and, in some ways, derivative of the genome project. The HGP can provide an invaluable resource for planning, designing and launching a research program focused on human diversity, but such a program would clearly not fit the central goals of the HGP. To divert any part of the HGP to address these issues now would be a great disservice to other areas of biology and medicine that are likewise dependent on the acquisition of the knowledge resources being derived from the genome project, and which could make similar claims to scarce sources of support. In addition, it must be noted that the more focused and timely the acquisition of the HGP data is, the more effective a human genetic diversity project will be. Thus, any diversion from the goals of the HGP would, in the long view, detract from diversity studies as well as others.

There are, then, substantive reasons for keeping diversity studies distinct from the HGP. Let me now turn for a moment to some considerations on the planning for such a project. First, I will discuss an issue that is distinct from the scientific questions. Before starting an initiative like this one, serious thought must be given to the impact on the population to be studied, especially in the area of ethical, legal, and social implications of the research itself and the information it could produce. These are serious issues. Privacy, discrimination and legal and social issues in the research locales are significant considerations in planning such a project, and must not be slighted. Of course, since all federal agencies have subscribed to a "Federal Policy for the Protection of Human Subjects" (10 CFR Part 745)—and part of this rule is that all human subjects, anywhere in the world, must be treated as they would be in this country—all subjects and all information deriving from this research would have to be treated with the great care specified in those rules. In addition, Third World countries may not feel that providing unlimited access to their "genetic resources" without assurances of various kinds, perhaps including compensation, is in their best interests. Also, how other research workers, unconnected with the immediate research but having access to the data, might make use of biological materials and information created by this project is a serious issue and must be considered. These difficult issues have not received much attention to date. It is clear to me that it would be irresponsible not to have these issues clearly addressed before such research ensues.

I have argued that the human diversity research should not be considered to be a part of the HGP. I have also pointed out that there are several societal obstacles that must be overcome to carry out such research in an effective and responsible manner. I would like to mention now that there are significant scientific issues that need to be considered further before this research project should be considered for significant support.

While the differences between two human genomes can be significant in what they could tell us about their respective genetic origins, the differences between two human genomes appear to be, at any one position, of the order of one part in one thousand. There are regions that are highly variable, changing rapidly so that only close relatives share the same sequence—the basis for DNA forensics. There are also much more slowly varying regions that are shared by most individuals with relatively distant genetic heritage. What this means is that very careful considerations must go into devising a plan for sampling a human population before it is likely to be useful. The genetic differences being looked for must be anticipated to some extent in order to plan the sampling properly. If one is searching for characteristic genetic differences that can mark historical and recent pre-historical populations, one must be very careful to understand the genetic markers one is studying, and tailor the sampling scheme to the genetics. It has also been suggested that very important information about medical genetics is to be gained by sampling diverse human populations. While this is certainly true in principle, it is also clear that this kind of information is unlikely to derive from a program that takes a few samples

from a widely diverse set of groups and individuals. The point of this discussion is simply that a great deal remains to be done to arrive at a robust plan for such a program if it is likely to be useful to science and to humanity.

Because of our interest in fostering the productive use of genome research in related areas, the DOE, as well as the NIH and the National Science Foundation, have provided some support to interested scientists to hold several planning meetings over the course of the past year concerning the study of human genetic diversity. It is reasonably clear to me that this planning process has not yet succeeded in devising an incisive set of goals and objectives, a convincing rationale, and a clear set of immediately achievable goals. This can clearly be achieved, but at the moment the plans remain immature. The important point, however, is that when the scientific questions and the ethical and social issues are properly dealt with in planning this research, the project should be considered only on its own merits, and certainly as distinct from the HGP. If this planning is done well, the project may well contribute substantially to understanding our biological heritage and our history—invaluable contributions to human knowledge. This represents another area that the HGP can contribute to by providing the fundamental resource for human diversity research.

Conclusion

The Human Genome Project, the historic, coordinated, international effort to map and sequence the genetic material of our species, is well underway and is making great progress. There are many reasons that this enterprise was undertaken; its ultimate utility should extend to every branch of biological science, medicine, and biotechnology. Profound questions about human biology, human origins, human development, and human disease, will become answerable. To accomplish this in a reasonable time period and with the anticipated efficiency, which results from adherence to a careful plan, we must focus our attention and our resources on the goals at hand. The beneficiaries of the HGP will be every branch of biology and medicine. I support the broad aims of those interested in human genetic diversity studies and encourage them to make use of the powerful tools, resources, and vast knowledge deriving from the Human Genome Project. However, these studies should compete for funding on their own merits and should not be a part of the Human Genome Project.

This concludes my prepared testimony. I would be happy to answer your questions.

Senator AKAKA. Dr. Nishimi?

TESTIMONY OF DR. ROBYN NISHIMI, SENIOR ASSOCIATE, BIOLOGICAL AND BEHAVIORAL SCIENCES PROGRAM, OFFICE OF TECHNOLOGY ASSESSMENT

Dr. NISHIMI. Thank you, Mr. Chairman.

I appreciate the opportunity to appear before you to discuss the issues raised by the Human Genome Diversity Project. My written testimony presents a general overview of this project and a broad identification of the technical, funding, and ethical, legal and social issues that it might raise. I will focus my oral testimony on summarizing matters related to funding and to ethical, legal and social implications.

The Human Genome Diversity Project, as you have heard, is a proposed and unfunded effort. A congressional role is important because, even at this early stage, it is obvious to OTA that the project is unlikely to move forward in any coordinated fashion without some Federal funding.

Such funding might derive from three sources: NSF, NIH, or DOE, singly or in collaboration. It appears to OTA, however, that prior to Federal funding, an analysis is necessary to determine what level of funding is needed, how much is currently available from reprogramming of existing NSF, NIH or DOE funds, and what new appropriations might be necessary.

Scientists proposing the project estimate its cost at about $25 million over 5 years. OTA has neither analyzed nor verified this estimate.

With respect to ethical, legal and social issues, I would like to highlight a few questions. First, many of the populations proposed for sample collection are in developing nations. What if certain data gathered prove commercially valuable? Developing countries have already expressed concerns about patent issues, such as those spotlighted by the Rio de Janeiro summit in the context of intellectual property protection for novel plant and animal material and the "Biodiversity Treaty." Not surprisingly, concern is heightened by the prospect that the substance now in question is human biological material from vulnerable populations.

Beyond the legal issues of intellectual property protection are also ethical and social questions, and I will just focus on a few.

For example, will any benefits of the research accrue to the research subjects? The issue of informed consent itself raises several questions. Informed consent expressed in whatever form is a minimum activity to demonstrate respect for a culture. Do Western notions of informed consent have any true relevance to some of the populations to be sampled? Yet, while Western notions of consent might differ drastically from those populations to be studied, the concept of respect surely persists.

Clearly, any research conducted with Federal funds will need to comply with current U.S. regulations governing human subjects research. Yet these rules may conflict with the practices, values or beliefs in other societies.

What about, for example, the U.S. practice of a written document? A written document is important because it endures as a record for the future, either as an instructive tool or for auditing purposes; yet written documents will be an anathema in some cultures.

Similarly, Federal regulations require that a project's purposes be explained. One of this project's goals is to elucidate information about the origins of the sample population and its relationship to other populations. Some cultures, again, however, have deeply-rooted beliefs about their origins and would find this goal offensive or insulting. Would it be ethical to emphasize certain goals—for example, identifying disease susceptibility—over others—for example, examining human origins—in order to merely facilitate consent and participation?

Another example is the issue of obtaining informed consent from every individual in a community, or can a local leader or other central authority speak for all members? Consider a situation, for example, of a local leader who agrees to sampling for all members of his population, but researchers who are preparing to draw the blood from a woman plainly see that she is distressed by the prospect. Do they proceed, or do they decide not to sample? Will this now incur the leader's wrath because he has lost face, since he had given his word that all would participate? Will the woman now be overtly punished, or if not overtly punished, will this stigmatize her?

Analyzing ethical considerations is especially critical because many of the populations that have been proposed for sampling are groups that historically have been vulnerable or exploited.

Mr. Chairman, as was my task, I have focused on identifying issues that the Human Genome Diversity Project raises. Nevertheless I think it is important to stress that despite the questions I have enumerated, OTA does not view this project either negatively or positively. Rather, I emphasize that OTA supports a thoughtful and deliberate discussion so that Congress can consider a full range of options. Clearly, a balance must be struck between our intellectual desire to pursue an exciting and interesting line of inquiry against both our responsibility to devote research resources in the most efficient and prioritized manner, and more importantly, our ethical obligation to respect and enhance the welfare of all people.

Thank you. I will be happy to answer any questions.

Senator AKAKA. Thank you very much, Dr. Nishimi.

PREPARED STATEMENT OF DR. ROBYN Y. NISHIMI

Mr. Chairman and Members of the Committee, I appreciate the opportunity to appear before you today to discuss the opportunities and concerns raised by the proposed Human Genome Diversity Project. Some of these issues are related to the Office of Technology Assessment's (OTA) ongoing study, "The Human Genome Project and Patenting DNA Sequences." This assessment, which I direct, is scheduled for delivery to the Technology Assessment Board in April 1994.

I would like to emphasize at the outset, however, that my remarks will focus on the Human Genome Diversity Project, which is distinct and separate from the Human Genome Project. My comments will be a general overview of the Human Genome Diversity Project and a broad identification of some of the issues that the project might raise. Currently, OTA is not conducting a full assessment of the Human Genome Diversity Project, as it did for the Human Genome Project in its 1988 report *Mapping Our Genes—The Genome Projects: How Big, How Fast?* Nevertheless, I anticipate that some of our current analyses for the "DNA patents" project—e.g., of intellectual property protection of DNA sequences and technology transfer issues in—will be pertinent to a few aspects of the Human Genome Diversity Project. As I will elaborate further, however, some issues raised by the Human Genome Diversity Project are beyond the scope of OTA's current study.

BACKGROUND

In humans, as in essentially all forms of life, deoxyribonucleic acid—DNA—contains the entire genetic blueprint for an individual. Today, scientists in the United States and abroad have undertaken the 15-year, $3 billion Human Genome Project. The result of this effort will be a single "reference" map of composite information that is essentially a Caucasian genome. Yet no two individuals, except identical twins, share the same DNA sequence. Furthermore, genetic diversity clearly exists among populations around the world. The Human Genome Diversity Project proposes a systematic examination of human DNA sequence variation by sampling 20–25 unrelated individuals from each of 400–500 populations of historical interest. It would be undertaken with the expectation that fundamental questions about the origins, settlement, and migration of humans could be examined. As well, the project could elucidate why some populations are more, or less, susceptible to certain diseases.

Discussion—and some early decisionmaking—on *if, when,* and *how* to undertake the Human Genome Diversity Project is a matter of some urgency: Several of the proposed research populations are literally becoming extinct.

FEDERAL FUNDING

The Human Genome Diversity Project is a *proposed* effort, for which its supporters currently seek funding, and I would like to emphasize the importance of a congressional role: Even at this early stage, it is obvious to OTA that the Human Genome Diversity Project is unlikely to move forward in any coordinated fashion with-

out some U.S. Federal funding. Scientists proposing the project estimate its cost at $23–25 million over five years; OTA has neither analyzed nor verified this estimate.

Federal funding for the Human Genome Diversity Project might derive from three sources—singly or in collaboration: The National Science Foundation (NSF), the National Institutes of Health (NIH), or the U.S. Department of Energy (DOE). NSF is appropriate because it is the primary Federal agency that funds anthropology research, and most—but not all—of the questions that might be addressed by the Human Genome Diversity Project are anthropological in nature. NIH and DOE are appropriate because together they fund the U.S. Human Genome Project, and thus are familiar with the latest developments in genetic research. OTA has performed no analyses on funding structure or outlays for the proposed Human Genome Diversity Project.

Since each of the agencies is represented at this hearing, I will not elaborate further on individual or joint capacities to fund the Human Genome Diversity Project. I note, however, that neither NIH nor DOE has included funding for the Human Genome Diversity Project in its current 5-year plans for the Human Genome Project. Hence, Congress would need either to appropriate additional funds to NIH and/or DOE for the Human Genome Diversity Project or instruct the agencies to redirect existing funds to the diversity project—at the expense of investigations funded under the auspices of the Human Genome Project. Further, I am not familiar with the details of funding for anthropological research through NSF and the extent to which NSF could fund this project—again solely or in cooperation with NIH and DOE.

In sum, it appears that if the Human Genome Diversity Project is to receive Federal funds, an analysis is required to determine what level of funds is needed to conduct the project, how much is currently available from reprogramming of NSF, NIH, or DOE funds, and what new appropriations might be necessary.

TECHNICAL ISSUES

The Human Genome Diversity Project presents several technical issues, which other witnesses can certainly address in greater detail. Briefly, some considerations include the following.

- Although preliminary work has been undertaken to identify the populations for DNA sampling, Dr. Richard Ward, an anthropologist involved in the initial effort raises concern that the criteria for identification vary among the groups, which were divided by geographic area. The issue is: What criteria should be employed to create a final sampling list to ensure that the project is consistently implemented? Then, does the current proposed list meet the criteria?
- The project proposes to collect blood samples from indigenous peoples. Certain blood cells will be transformed into what scientists call "immortal cell lines." This process must occur within 72 hours of collection, but offers the advantage of largely preserving the genetic heritage of an individual in perpetuity. Nevertheless, certain interesting genes—in particular those involved in immunity to disease—will not be amenable for analysis alter transformation; other DNA changes, currently unknown to scientists, might also be associated with transformation. Thus, besides collecting blood samples, what other biological material should be collected for DNA analysis (e.g., hair, cheek swabs)? From whom should this material be collected? Just the 25 individuals from whom blood has been drawn? Since interesting information can be derived from DNA samples alone, how many additional human research subjects are proposed for supplemental sampling?
- What biological material will be stored? How will it be stored? Who will manage access? Who will oversee quality control and quality assurance? Immortal cell lines require ongoing (in perpetuity) upkeep: Who will pay for these costs? What kind of databases will be used to manage the storage and retrieval of biological material? What about informatics for the information generated from the biological material? As with the biological material, who will manage the information quality, access, and upkeep?
- What minimum set of genetic markers should be analyzed for all samples? In what priority will samples be analyzed once a common set of markers has been identified?

Although these are important questions, OTA's forthcoming report on patenting DNA addresses none of them.

ETHICAL, LEGAL, AND SOCIAL ISSUES

The proposed Human Genome Diversity Project raises several ethical, legal, and social issues, and it is impossible to fully discuss them within the confines of my testimony. It is critical, however, that these issues be addressed.

For example, many of the populations proposed for sample collection are in developing nations. What if certain data gathered under the Human Genome Diversity Project prove commercially valuable? The United States must be sensitive to the concerns of developing nations, while simultaneously preserving legitimate interests of U.S. companies to pursue commercial development and intellectual property protection of "biotechnological" products. Developing countries have already expressed concerns about the Human Genome Project and the general issues surrounding DNA patents. These concerns are similar to those that were spotlighted at the Rio de Janeiro summit last year in the context of intellectual property protection of novel plant and animal biological material and the "Biodiversity Treaty." Not surprisingly, however, they are heightened by the prospect that the substance now in question is human biological material from vulnerable populations. As Chaim Sheba of Israel has put it, "You have taken our gold and diamonds; now you are taking our genes."

Beyond legal issues of intellectual property protection are several ethical and social questions.

- Will any benefits of the research accrue to the research subjects?
- What are the risks—particularly social risks such as stigmatization—to research subjects who participate? To those who decline to participate?
- What about compensation—monetary or otherwise—for research subjects? What if a country seeks payment in return for the collection of biological samples from its citizens? What if a local leader demands compensation? If compensating individuals is an option, what happens to people not "chosen" as research subjects? What benefits realistically can be offered to participating communities, especially if they are small and isolated?
- What about confidentiality of the samples collected? What limits are necessary? Who decides? Who will be responsible for ensuring confidentiality?
- In addition to collecting samples for the Human Genome Diversity Project, what about testing blood samples for disease? In particular, special attention to HIV infection is an issue—especially with blood samples from Africa, for example. Should samples be tested so that researchers can exercise greater caution when handling certain samples? Should the samples be tested anonymously to determine infection rates in these populations? Is it ethical to test samples anonymously? If it is unethical, then how does a demand to link a sample with an individual for purposes of disease identification balance against the importance of confidentiality for other purposes? If anonymous testing is viewed favorably and HIV testing is deemed desirable for the safety of scientists or to analyze HIV infection in indigenous populations, how will pretest counseling for HIV be handled on top of informed consent for the Human Genome Diversity Project, generally?
- The issue of informed consent raises its own questions: What constitutes meaningful informed consent in non-Western cultures? In fact, do Western notions of informed consent have any true relevance to some of the populations to be sampled? Nevertheless, informed consent, expressed in whatever form, is a minimum activity to demonstrate respect for a culture. And while Western notions of consent might differ drastically from those of the populations to be studied, the concept of respect surely persists.

  Nevertheless, any research conducted with Federal funds will need to comply with current U.S. regulations governing human subjects research, in addition to adhering to any local governmental rules. U.S. regulations lay out eight specific informed consent requirements, and I will mention at least three that may result in conflict between U.S. regulations and practices, values, or beliefs in other societies: a requirement for a written document, a clear explanation of the purposes of the research, and individual consent.

  A written document is important because it endures as a record for the future—either as an instructive tool or for auditing purposes. Yet written documents will be an anathema in some populations. How can this be reconciled with Federal regulations?

  Similarly, Federal regulations require that a project's purposes be explained. One of the goals of the Human Genome Diversity Project is to elucidate information about the origin of the sample population, as well as its relationship to other populations. On face value, this might easily be put

in terms understandable to all cultures. However, some cultures have deeply rooted beliefs about their origins and would find this goal of the project insulting or offensive. What if this jeopardizes efforts to obtain samples from key communities? Would it be ethical to emphasize certain goals—e.g., identifying disease susceptibility—over others—e.g., examining human origins—in order to facilitate consent and participation? Who decides?

Must informed consent be obtained from every individual in a community? Can local leaders and other central authorities decide for all members of a community? Consider for example, a situation where a local leader speaks for all members of his group. What if he agrees to the sampling, but researchers who are preparing to draw blood from a woman plainly see that the woman is distressed by the prospect? Do they proceed? Do they decide not to sample her blood? Will this now incur the wrath of her leader, who has "lost face" because he had given his word that all would participate? Will she be punished overtly? If not overtly, will this stigmatize her? Should the investigators try to gain her individual consent? Is any consent she might then give truly consent, or has it been coerced?

- If protocols vary (e.g., regarding compensation or consent) from culture to culture, as might be expected, who will arbitrate what's necessary to ensure the protection of the human research subjects? Will population-by-population approval be necessary—especially if U.S. funds are used? If individual protocol review is deemed cumbersome by researchers, is there a reasonable expectation that all contingencies can be identified prior to embarking on sample collection?
- What about issues of genetic discrimination? What is the best mechanism to minimize misuse and misinterpretation of the gathered data? In particular, since genetic differences are the proposed project's focus, concerns are raised about information being used to support notions of superiority of one group over another. Dr. Diane Paul, who has studied the social history of genetics, believes that the Human Genome Diversity Project is likely to reinforce conventional views of race and ethnicity.

I have elaborated on ethical, legal, and social issues at some length, but by no means in an exhaustive fashion. I reiterate that analyzing these considerations is especially critical because many of the populations that have been proposed for sampling are groups that historically have been vulnerable or exploited. And while anthropologists have dealt with some of these questions in their research for several years, the scale of the proposed Human Genome Diversity Project changes the dynamics, as does the current international political context.

The forthcoming OTA report on intellectual property protection for human DNA sequences will examine the question of patenting DNA, per se. However, it will not directly address the issue of property rights that might be sought from samples such as those that would be gathered under the Human Genome Diversity Project. It also will not analyze the ethical issues I have just touched on.

THE OUTLOOK FOR THE HUMAN GENOME DIVERSITY PROJECT

Mr. Chairman, as was my task, I have focused on identifying issues that the Human Genome Diversity Project raises and have not offered specific options to address them. Nevertheless, I think it is important in concluding my testimony to stress that, despite the many questions I have enumerated, OTA does not view the Human Genome Diversity Project either negatively or positively. Rather, I emphasize that OTA supports a thoughtful and deliberate discussion involving a broad spectrum of international perspectives, so that Congress can consider a full range of options. Clearly, a balance must be struck between our intellectual desire to pursue an interesting, exciting line of inquiry—one that may yield information about human origins, as well as medically important data—against both our responsibility to devote research resources in the most efficient and prioritized manner and, more importantly, our ethical obligation to respect and enhance the welfare of all peoples.

Mr. Chairman, thank you for the invitation to discuss this important topic. I will be happy to answer any questions you or Members of the Committee might have.

Senator AKAKA. I thank all of you for your testimonies. They will be helpful. I have a number of questions for you and would like your responses.

I think you know that the reason we are having this hearing is to get into the record some of the important data that can help us in the future as we continue to move the program and to try to get more funds than we have had in the past.

I would like to focus this hearing on the Human Genome Diversity Project, and its relation to the much larger Human Genome Initiative.

The study of the genetics of ethnic and aboriginal populations promises to answer many social and anthropological questions about mankind. That is what makes it exciting. I do not question the importance of that information. I believe, however, that Federal funding for such a project must be based on the potential for yielding tangible benefits for medical science.

Let me direct my first questions to the DOE and NIH witnesses. In your testimony, you mentioned some of the benefits that might come from the study of diverse genomes. In what ways could such a project complement the current Human Genome Initiative?

Dr. Galas, since you mentioned the relationship, I'll call on you first.

Dr. GALAS. I think it is very important to understand that the human genome, as we view it from the point of view of the Human Genome Project, is a very complex object because it represents, if you will, a map of all of our genetic blueprints, which differ from individual to individual, as the next panel will discuss in a little more detail.

I think that the precise benefits to the Genome Project of the Diversity Project depends on how you frame very specific questions. I don't think there is a general answer to that, because if you wish to answer certain questions about variations in certain genetic loci, or a certain type of variation in the genome between individual chromosomes or some such well-framed question, then you can answer the question that you ask—what is the potential complementary benefit.

It is certainly true that having diverse, different genomes to look at can be of potential interest to doing the very basic things, mapping the humane genome. But unless those questions are framed in a very narrow and very critical way, I don't think that those questions can be answered.

Dr. COLLINS. Can I address that as well?

Senator AKAKA. Please.

Dr. COLLINS. The Human Genome Project has defined its goals carefully in order to be sure that the job gets done, and the project is aiming to develop maps, both genetic and physical, and to sequence the human genome. The Human Genome Project does not intend itself to apply that information to all of the medical conditions that would be benefitted by such an approach. So, for instance, in the search that is going on right now for the breast cancer gene, to a major extent the reagents and the maps that are being used for this process are products of the Human Genome Project. The utilization of those maps, however, to target that very specific breast cancer gene on Chromosome 17 is an effort that is also supported by other parts of the NIH (especially NCI), and should be.

The Genome Project is there to build what we would loosely call infrastructure, to make it possible for projects like this to go forward. These would have been unthinkable without the rich array of resources that the project is now providing.

Similarly, I think the Diversity Project has goals in common with the Genome Project, and there are areas of significant overlap particularly in technology development. The attempt to look at the DNA sequence of many different individuals, which is what the Diversity Project needs to do, is similar to what needs to be done when you are looking at families in which a particular disorder is occurring, and you are trying to figure out what genes are responsible. So there are shared areas of interest which I think could be exploited.

But in current terms, it would be difficult for the Genome Project to expand its umbrella to take on other more diverse efforts like the Diversity Project without expanding its funding.

As far as the medical consequences of the Diversity Project, this is a difficult concept to think through. To what extent would having blood samples from 400 different ethnic populations help you understand why certain populations are predisposed to some diseases and others are not? Part of the answer comes down to what those genes are going to look like.

Just to use an analogy, consider a group of people wearing different colored sweaters. Let us suppose it is the people with the red sweaters who are somewhat more predisposed to getting a disorder, say diabetes, than the people who have green sweaters or blue sweaters or yellow sweaters. By comparing ethnic groups which have different proportions of red sweater wearers you might deduce the correlation. But that only works if there are a lot of people who wear red sweaters. If you are only looking at 25 people in each group, you need to have a fairly frequent finding in order to draw a conclusion.

If, on the other hand, you are looking for a relatively rare gene—say, a red sweater with a blue collar—then 25 people may not include one of those, and you may therefore not be able to draw any conclusions about why this particular population has a higher frequency of a particular disease.

I do think we will learn things indirectly in either event about medical conditions, simply by finding out relatedness of populations. If you find that this population has a high frequency of diabetes, and another one does also, and then it turns out that they are related to each other, that implies to you that this is probably genetically programmed, especially if their environmental situation is very different; but it is a bit of an indirect argument.

So I think it is really not easy to say right now what the solid medical consequences of the Diversity Project might be. They will occur. I am not exactly sure what they will look like and how quickly they will come along.

Senator AKAKA. The primary mission of the diversity study would be the collection and preservation of genetic information from several hundred—you point out 400—ethnic and aboriginal groups that are near extinction. I use the word "extinction" loosely, and I am simply emphasizing that many of these populations are on the verge of losing their unique genetic identities.

Do you feel there is merit in sampling ethnic and aboriginal populations while it is still possible, so that the information can be preserved for future investigators? I am still asking questions of DOE and NIH.

Dr. COLLINS. I am happy to answer that question. Yes, I believe there is great merit. I do think the urgency is a compelling argument, that populations are not going to be there indefinitely for us to decide at some future date we would like to sample them.

So I do believe there is merit in the proposal. Obviously, that has to be couched with all the concerns about the ethical issues, which I think can be dealt with, but probably should be dealt with very soon.

The Diversity Project in its first three workshops has addressed those issues and has, I think, made a good initial review. I think it is very important that there be broad discussion of those issues and input from some of the countries that would be involved to be sure that the necessary safeguards are in place.

Senator AKAKA. Dr. Galas?

Dr. GALAS. I would agree, Senator, that there is certainly merit. So I guess I would agree in principal with what Francis said. It seems to me, however, that the devil is in the details, as we say, and that precisely what questions are being asked, or what we want to find out, will determine how the sampling is done, what sort of planning is carried out, and whether or not it is just sampling, or whether these people are to be looked at with respect to their genetic phenotypes as well. If you want to understand something about the medical implications, you have to know something about the medical condition of the person from which these came.

So these are questions that it seems to me need to be answered very precisely and very carefully before one can quantify the merit, however great it may be.

Senator AKAKA. It is always a pleasure to work with the talented scientists and engineers at OTA——

Dr. NISHIMI. That's a dangerous lead-in.

[Laughter.]

Senator AKAKA. You certainly make our job easier, we think. Thank you for being with us today.

Dr. NISHIMI. Thank you.

Senator AKAKA. We are all aware of the wonderful new treatments that might develop from the Human Genome Initiative. Have individuals from your agency studied the potential benefits to medical science that the Diversity Project might offer? If so, could you please detail for us what those benefits would be?

Dr. NISHIMI. OTA has not looked at the possibility of medical benefits from the Human Genome Diversity Project per se. We have looked at the Human Genome Project, but not the Diversity Project.

Senator AKAKA. What are the advantages of collecting DNA samples for diverse populations?

Dr. NISHIMI. I think you have heard some of the benefits that might accrue from this project. I describe a few of those in my prepared statement—certainly, anthropological questions of human settlement patterns, migrations, and the like, and also the possibility of certain medical benefits. But OTA hasn't done a detailed analysis of either approach, so I am afraid I cannot offer you more than just generalities.

Senator AKAKA. Do you feel that the Human Genome Diversity Project complements the goals and objectives of the larger Human

Genome Initiative, and how might the two projects work together to help ensure the overall success of both?

Dr. NISHIMI. Certainly, there are common elements, as Dr. Galas described, in terms of the Human Genome Diversity Project learning from some of the Human Genome Project's technologies and being able to take advantage of those. I wouldn't like to offer, however, any analysis on possible funding structure relationships between NSF, NIH and DOE, or whether one agency should be the lead agency, because OTA, again, has not analyzed that question.

Senator AKAKA. The National Science Foundation has been a longstanding supporter of both hard and soft sciences, including anthropology and archaeology. I know that NSF supports several NSF Fellows studying anthropology. I am committed to the research opportunities offered by NSF, and I believe the educational benefits to the many supported students are a good investment by the United States. Thank you for being here today.

Has NSF funded projects similar in scope or purpose to the Human Genome Diversity Project? If so, could you briefly outline the merits on which NSF based their decisions?

Dr. MARRETT. NSF has funded any number of projects that bear on this particular proposal. In fact, I appreciate this opportunity to comment on some of the observations that OTA has made in terms of the ethical and other questions.

A number of those questions are in fact questions that are found in research, whether it happens to be on genetic diversity or not. One of the reasons why we have thought it very useful to combine the discussion of genetics with cultural and physical anthropology is that growing out of the research experiences particularly of cultural anthropologists has been an understanding of how to handle sensitive populations, sensitive matters.

If one looks at the second of the workshops held, for example, that workshop said it would be extremely important to integrate, to sample in ways in which one could identify those populations from which we already had information about access, about the other sensitivities, populations on which we would have other kinds of medical information. So I would stress that what we are trying to do is to indicate that we would not be starting from ground zero; it would be building on the body of knowledge that is accumulated about particular populations and particular procedures to make sure that there is that sensitivity to the sets of issues that clearly must arise.

Senator AKAKA. Are you comfortable, Dr. Marrett, with the methodology of the Diversity Project and its dedication to minimizing the misuse of genetic information?

Dr. MARRETT. Let me start with what I would hope we would share with all of my colleagues from all of the Federal agencies, that we see our task first and foremost as promoting science, and in our case, science and engineering, for the advancement of humankind.

Having taken that position, then, it would be in fact inconsistent to support activities that seemed harmful to humankind. So with that as our first basis, we would say we would pay close attention to any matters that might in fact jeopardize the human population.

Beyond that, I would certainly support what my colleagues here on the panel have said—that there are matters still to be worked out. There are numbers of issues having to do with procedures to be developed. That is why, again, we continue to stress this as a proposal and not as a full-fledged project. Why we would see a particular need and importance of linking in with the human genome organization is because of the expertise, the experience that has developed, not only on the technological side, but on the fact that that organization has been grappling with in very serious ways the questions of ethical, legal, other matters.

Hence, if you ask have all the issues been resolved so far, no, they have not been; they would be a part of the continued development that would be necessary if this proposal is to become a full-fledged project.

Senator AKAKA. I thank this panel for your testimony. I could detect from each of you the concern about being very careful with the information that will be collected, and I hope that as we move along in positive directions, we can work on this. As new ideas come forth, in new areas, there is always the legal aspect to consider. We need to keep an eye on this as we move down the road.

I want to say thank you very much for your testimony, Panel 1, and to tell you again that you have been very helpful.

Senator AKAKA. Let me now introduce Panel 2 and call to the table Dr. Luca Cavalli-Sforza, Professor Emeritus of Genetics, Department of Genetics, School of Medicine, Stanford University; and Dr. Mary-Claire King, Professor of Genetics and Epidemiology at the University of California at Berkeley.

I welcome you here today and look forward to your testimonies. As soon as you are ready, you are welcome to begin, Dr. Cavalli-Sforza.

TESTIMONY OF DR. LUIGI LUCA CAVALLI-SFORZA, PROFESSOR EMERITUS OF GENETICS, DEPARTMENT OF GENETICS, SCHOOL OF MEDICINE, STANFORD UNIVERSITY

Dr. CAVALLI-SFORZA. I am very pleased and honored to be able to testify on this important subject. I have spent much of my career in this field and now see the coming together of an urgent need for action and an increasing technological ability to take that action.

In my spoken submission, I will briefly lay out what the HGD Project will do. It is made up of three main parts—first, the collection of samples; second, their transformation and preservation, and third, their analysis.

In previous workshops organized thanks to help from Federal foundations, we have started discussing with the experts of various disciplines statistical, anthropological, molecular, ethical and legal aspects of our work, and we will continue to refine our plans as necessary.

We aim to collect a general sample of about 10,000 individuals representing the indigenous world population. This is the sample that Professor Francis Collins referred to before, but we also have planned to collect at the same time a larger sample of perhaps about 100,000 individuals, maybe 10 times as many, just for answering those questions that Professor Collins was rightly worried about that we might not be able to answer with the 10,000 sample.

The 10,000 individuals, which are the 400 populations with 25 individuals each, will be treated in a more expensive way (by Epstein Barr virus transformation) to make them available forever, so to say; but the extra, 10 times as many, individuals will not be equally transformed, so it will be a much cheaper sample to collect and store, and they will answer all those questions that may come up in the meanwhile, especially problems of medical genetics also for rare genes. It is a much cheaper thing to collect the 100,000 individuals sample once we have anthropologists in the field for collecting the smaller, more expensive samples.

After the samples are brought to laboratories, they will be transformed by a technique developed for the purpose of obtaining indefinite amounts of DNA, a procedure largely employed since the beginning of the eighties in research on medical genetics.

From 1966 to 1976, I have studied about 1,500 African pygmies in central Africa, and I returned in 1984 and 1985 to collect new samples. In collaboration with Yale University, these samples were transformed according to this new technique and were the first in a pilot project which now includes some 20 populations from all parts of the world. This gives us a guarantee that the project can be made. The samples from this collection have already been incorporated in the Standard Repository of Cell Cultures set up by the National Institutes of Health General Medical Sciences at Camden, New Jersey and are being increasingly used by scientists.

We have carried out extensive studies of the data collected on the populations prior to the DNA era and on those of the pilot project and believe that once new populations will be collected, their analysis should begin soon, as it could influence ways in which the project will be continued. In other words, we plan to learn from experience as soon as we start collecting samples.

New technology being developed for the Human Genome Project will soon become available, we hope, and will certainly be extremely useful. It is our desire to keep in close contact with that project at all levels, from molecular to data analysis, and to maintenance of databases. Cell lines, DNA results, are not intended as a private reserve, but must be available to every scientist. The cost for access to these resources should be assessed in the best interest of science and society.

What can we expect to find? There are historical, anthropological, medical implications. Professor Mary-Claire King will speak about the last ones. In a book of mine on the history and geography of human gene, which is an analysis of what is already known, and which is to appear before the end of this year, there is a summary of conclusions obtained from the analysis of thousands of publications that have appeared so far, mostly generated in the pre-DNA era. The transition to DNA technology has greatly amplified the power of the genetic approach but has not contradicted the results that we have already obtained. These remain the most powerful scientific tool we have against racism, as summarized in my written testimony.

Conclusions of the most recent studies have raised considerable interest among scientists of many disciplines, also outside the field of genetics, ranging from history and archaeology to anthropology and linguistics.

I have been amazed especially by the excitement among laymen. I have been overwhelmed by the volume of correspondence I have received. It is clear that this work will help educators to spread public knowledge of science and create new interest for it. It is a beautiful example of how exciting science can be. It will also generate more understanding of genetic differences between individuals and between groups, which are very small compared to those between individuals of one group, as well as of the rights of indigenous people. One should not forget that public awareness can be very beneficial to indigenous populations, who continue to be abused even today in many countries. But it is essential that the information be accurate, or it could be counterproductive.

Whether for science or for humanitarian purposes, it is important that the field work be carried out by scientifically-trained anthropologists. This is a major departure that we propose with respect to previous collection of these data.

We must not forget that it is urgent to undertake this work before economic development completely destroys or confuses the cultures and identities of those populations. Just to give a few examples, in Tierra del Fuego, there are left two or three descendants of the Ona tribe, if any, and perhaps two dozen of the Alacaluf nearby. These are people who were there at the time Charles Darwin visited the area. Only a few dozen individuals from four Negrito groups survive in the Andaman Islands and have very low fertility. There are extremely few children. How many, if any, will there be at the turn of the century? We cannot afford to wait.

Thank you.

Senator AKAKA. Thank you very much, Dr. Cavalli-Sforza.

PREPARED STATEMENT OF DR. L. LUCA CAVALLI-SFORZA

1. The general aims of the HGD Project

The Human Genome Diversity (HGD) Project aims to survey the genetic diversity of living humans. It proposes to do so by sampling primarily indigenous populations, whose origin is usually better known than that of individuals who emigrated in the past centuries to new places. These samples will come from an adequate number of individuals representing the human species and will be used to form a repository ensuring unlimited survival of their genetic material, DNA, for future study. Special, but not exclusive, attention will be paid to peoples of historical interest who are likely to disappear through physical extinction or, more frequently, dispersal and loss of identity. These events are taking place at an alarming rate, paralleling the rates of economic development and habitat destruction that cause them. The Project is therefore urgently needed in order to avoid the irreversible loss of precious human genetic information.

2. Storage of cell lines

The HGD Project makes use of the existence of modern techniques of cell conservation, which allow us to keep certain cells of an organism alive and capable of multiplying, thus generating potentially unlimited amounts of the organism's DNA. Almost no DNA deterioration is experienced under these conditions. For this purpose, fresh samples of blood must be collected and transported shortly after collection to appropriately equipped laboratories, where certain blood cells (B lymphocytes) are transformed by Epstein-Barr virus. The cells are then grown and stored on liquid nitrogen, where they conserve indefinitely their viability and capacity to multiply. This is a well-tested, standard procedure. There has also been a pilot project, which has accumulated a small number of samples of individuals from indigenous populations of various continents for research use.

3. Choice of populations and other technical problems

We have held three workshops with support from the National Institutes of Health (NIH), the National Science Foundation (NSF), and the Department of Energy (DOE)—in July 1992 at Stanford University, October 1992 at Pennsylvania

State University, and February 1993 at the NIH campus in Bethesda. These workshops have helped us form a preliminary opinion of the major problems at the statistical, anthropological, molecular, bioethical, and social levels. Final proceedings of these workshops have been or are being prepared, and we are in the process of forming working groups to continue the analysis. Appropriate choices of populations and individuals are an essential part of the Project, and to this end we have enlisted the cooperation of an international array of anthropologists. With their help, we have established a preliminary list of about 400 to 600 populations to be sampled. On the basis of discussions with population geneticists and statisticians, we plan to generate transformed cell lines from about 25 unrelated individuals in each population. Anthropologists already working with the population generally will obtain the blood samples for the preparation of transformed cell lines. In addition, we expect the anthropologists to collect blood, saliva, or hair samples from at least ten times as many individuals in the same and related, neighboring populations. These latter samples will not be subjected to transformation, but will be stored for special studies requiring larger numbers of individuals for which smaller, limited amounts of DNA are sufficient, and for which methods of DNA amplification that are less sophisticated, and less expensive, than Epstein-Barr virus transformation can be employed.

4. Extent of diversity of the human genome

It would be wrong to think that the human genome exists in a single copy, repeated again and again in every individual. The opposite is true; one can say that there are as many human genomes as there are humans, and the potential variety of human genomes is on the order of one followed by one million zeros, an extremely high number. It is very unlikely that any living human beings have exactly the same DNA. Even identical twins differ from each other, because after duplication of the fertilized egg giving origin to both of them, a dozen or more fresh mutations are expected to occur.

Prompted by the enormous magnitude of the task of analyzing the whole human genome, the Human Genome Project was started, in its simplest form, to study a single copy of the genome (not necessarily all coming from the same individual). Our lab is involved in research on two genetic diseases. We are well aware of how useful it would be and how much time we would save if we had the preliminary information on the DNA sequence in even one individual.

It is true, however, that our lab and others will gain a number of important benefits from a better understanding of individual variation. An important step forward can be made in the next 5 years with an expenditure that is a very small fraction of the annual cost of the Human Genome Project per year. The time is now ripe for starting an adequate analysis of diversity; and the optimal procedure is to begin it as soon as possible in concert with the sampling of disappearing indigenous world populations.

5. Why study individual diversity?

It is easier to appreciate the general importance of studying individual DNA variation if one considers that the real purpose of studying the sequence—the structure of the human genome—is to infer the function of its parts. This constitutes a considerable challenge. We cannot today accurately predict the function of a DNA segment by simply determining its sequence; we can only make some informed guesses. The study of variation of a DNA segment among individuals can, however, contribute in an important way to understanding its function. In general, genes that are basic to the life of the cell or of the individual vary very little from one individual to the other and only in specific ways, which affect only trivially their function. Lack of variation, or its minimal presence, indicates an "important" gene; changes in it are likely to determine serious abnormalities. For this reason, the study of genetic diseases can contribute in a major way to understanding the function of a gene. In fact, the identification of a gene responsible for a disease usually requires detection of a clear association between changes in its structure and specific pathological phenomena. More generally, the function of a gene is understood by making with assumptions about the specific function influenced by the gene together with a study of the variation among individuals of the DNA structure. Although some of this work can be carried out in test tubes, knowledge of individual diversity remains essential to truly understanding the ways genes function.

6. Some general guidelines for the study of diversity

Five years after the beginning of the Human Genome Project, the task of sequencing a single genome still appears formidable, and it is thus far premature to think that one could, or would want to, completely sequence even just two genomes, which might seem the minimum necessary for learning about individual diversity. But

there are rapid technological advances that should ensure progress in a few years, and what appears today an impossible task may well be far less difficult a few years from now. It is also certain that the technologies being developed for study of the human genome are ideally suited for the analysis of individual diversity, and that there will be considerable reciprocal fertilization at the technological level between the developments in the mainstream Human Genome Project and the Human Genome Diversity Project. And, after all, the applications of the HGD Project to medical genetics, which are one of the main justifications for the Project, are merely specific instances of individual variation. At least initially, however, it is neither feasible nor necessary to study individual variation for the whole human genome.

Rather than studying all individual variation in the entire human genome, we need to set up methods and learn rules that allow us to predict the individual variation to be expected in a given segment with minimal effort. This can be done by studying individual variation in a sample of DNA segments; this sample need not be large. I mentioned the rough guiding principle that the more important the function of a DNA segment, the less variation is seen in that segment among individuals. But this principle is not general enough. There exist DNA regions for which individual diversity is a biological necessity. We study the HLA genes, for instance, because of their importance to the success of organ transplantation and to the predisposition to certain autoimmune diseases. One could say that these HLA genes help a body know what is part of it. Here, extreme individual variation is essential for the HLA complex to work.

Because of its cost, the analysis of DNA samples from the world's population for the purpose of studying individual diversity can be extended by present methods only to a small fraction of the whole human genome. But the study should be planned so as to make that fraction as informative as possible. Let us assume that we want to analyze 10,000 DNA segments of a convenient length in different parts of the genome. They should be chosen to represent various categories of genes and DNA segments, which we can already identify to some extent. Although, ideally, they would be studied for all of the roughly 10,000 individuals from whom cell lines will be prepared, in Its early stages the analysis should be restricted to a smaller fraction of individuals (for example, a well-balanced mini-sample of 100 to 1,000 individuals). The mini-sample could be chosen in such a way as to represent the world through a hierarchical stratification of the 10,000 individuals. It would thus be easy to evaluate the usefulness of extending the survey of individual variation for a particular segment to all the 10,000 cell lines, or even to the 100,000 or more DNA samples from the non-transformed collection. Only those DNA segments that show sufficient variation would be eligible for study on the larger samples. The majority of them would not have to be sequenced, just tested with cheaper methods allowing the recognition of individual "polymorphic" (variable) sites. The collection of indigenous populations could be organized so that after the first year, it would constitute a "mini-sample" of individuals and populations adequate for coverage of the whole world. This mini-sample then could be employed for the initial screening of variation.

A very small collection of cell lines from all over the world is already in existence and is being used in preliminary observations, thanks to a pilot project set up in the last few years by the writer in collaboration with Professors Ken and Judy Kidd of the Genetics Department of Yale University. Samples of this collection are available to all research workers through the Camden Repository of New Jersey. But the mini-sample to be used for initial assays should be a better representative of the world than the one currently available.

Even with analysis of individual variation limited to a small fraction of the human genome, a general picture can be obtained that will help direct further investigations. We currently know very little of individual variation at the DNA level, but we are aware of the existence of some segments that are much more variable than others. We are bound to find more with a systematic search. The variation in these segments seems to have a precise biological meaning, and it will be important from a biological point of view to establish whether the high variation is due to high rates of mutation, natural selection, or other mechanisms. In my view, the search for highly variable DNA segments should have the highest priority because these segments are the most informative, but it will take time to detect them. A first phase of extensive preliminary investigations must therefore be dedicated to choosing appropriate, representative DNA segments to be studied. Neither in this initial phase nor in later ones will it be necessary to use full sequencing methods on all individual samples for each DNA segment; cheaper and faster methods would be preferable at least for initial screening, especially in regions in which variation is usually low.

It is my belief that research workers engaged in genome research will in due time find an inexhaustible source of important new biological problems in the study of individual variation. The HGD Project will take great advantage of, and also contribute in an important way to, the current trends toward automation of DNA sequencing and, especially, toward testing individual genetic variation of medical importance. Our rapidly increasing arsenal of research robots can speed up considerably the testing of large numbers of individuals for specific DNA segments, but we first need to learn more about the extent, location, and causes of individual variation, and more about optimal strategies for studying diversity.

7. The need to study variation throughout the world

Individual variation must be studied on a world sample, and not on a limited sample of individuals collected locally, as has been the case so far, because there exists intergroup as well as intragroup variation. We have recently demonstrated that limiting the study of variation to Caucasoids, as has been done in almost all the investigations of DNA variation, has introduced a serious bias in the evolutionary interpretation of the data. Moreover, the full power of the study of individual variation would not be harnessed if the study were limited to Caucasoids. There is considerable ethnic variation in genetic disease and predisposition to disease, which is important in planning health surveys, providing dietary advice, searching for donors for transplants, and so on.

It is thus necessary to explore the whole human species with a well-designed sampling scheme. At the same time, it is necessary to avoid the dispersion of effort that has characterized most work carried out in the pre-DNA era, in which different research workers studied haphazard collections of variable genetic characters on arbitrarily chosen populations. As a result of this lack of coordination of the earlier work, today one can obtain only an extremely patchy set of data from the published work. The information that can be obtained is a very small fraction of what it would have been if the immense effort of the thousands of research workers who volunteered their effort had proceeded in a more organized manner. A collection of cell lines made available to research workers in a central repository can help to avoid this waste by preserving a limitless supply of particular DNA samples to be analyzed. It will still be necessary, however, to ensure that some of the testing effort be made in a systematic way. By making available to researchers, at especially low cost, a set of cell lines chosen so as to be a representative world sample, the Project could help ensure a more balanced and informative effort.

I will concentrate here on giving a short account of scientific benefits that this endeavor can generate for the study of the history of human differentiation and evolution. We had already accumulated substantial knowledge on the genetic variation of human populations before the DNA era. But this information suffers from many uncertainties and from having been almost entirely unplanned and being therefore full of gaps, which make its analysis difficult and inefficient. Even so, we now know that the study of living populations allows the reconstruction of important aspects of human history. These reconstructions already have helped to explain archaeological, historical, and linguistic data. Striking conclusions have thus been reached, which have received wide interest, but which need confirmation and extension. Modern methods of studying individual diversity using DNA have much more power than older methods and can examine a much greater range of genetic variation (potentially all). They are much more efficient and precise and can be more easily automated than the older techniques; considerable efforts in this direction are already under way. Increasing the number of populations and the number of genes (DNA segments) studied thus far will lead to enormous increases in the variety of historical problems one can approach. Moreover, the analysis of DNA allows us to obtain "fossil" information, previously unattainable, from mummies, bones, and other ancient material. For the full use of fossil information, however, one needs to have adequate data on living populations from the same and neighboring areas, with which the fossil data can be compared.

8. Bioethical and social problems

Perhaps the most substantial question that arose in our workshop on bioethical and social problems is whether the study of intergroup genetic difference might foster racism. I am persuaded that it will have the opposite effect. It is difficult to do justice to this important subject in a short time, but more extensive analyses can be found elsewhere. It might be useful to begin by a definition of racism, which in my view is the belief that differences in the success of different peoples or races are due to their inherent biological superiority or inferiority. The idea is that this superiority is genetically determined and, as such, is inherited and impossible to change (notoriously, the superior race is almost always one's own). The importance of sociocultural inheritance in determining success (whether economic, political, or

military) is usually forgotten. All empires are of short duration, suggesting immediately that the background to success is socioeconomic and cultural; it cannot be genetic, because genetic processes are much slower, and very little genetic change can take place in the few centuries during which empires rise and fall. This is a very simple but powerful consideration that is not, however, heard frequently, simply because knowledge of genetic processes is confined to a few specialists.

Another consideration about racism is that genetic differences between human groups are small. We already have considerable information from a study of genetic markers with methods of the pre-DNA era. We have still-limited information on intergroup diversity using data obtained directly from DNA, since the relevant techniques have become available only in the last decade, but we already know enough to be confident that the picture thus obtained is very similar to the older one, but much more clear. The basic conclusion from the study of differences among groups is that they are small compared with the differences within the groups themselves. The aspiration to "race purity" of classical racism is absurd. A village or a small tribe will show almost the same extent of genetic variation among individuals as will the whole world. Only human populations of very small islands that have been subjected for a long time to very close inbreeding (marriage among close relatives) show a moderate increase in genetic homogeneity. This increase is often accompanied by infertility, as would be expected from the conclusions of countless animal (and to some extent plant) breeding experiments.

The statement that genetic differences among races are very small may seem to fly in the face of evidence known to everybody. If we look at skin, eye, or hair color, or the facial and body morphology of people who originated on different continents, we can usually predict their ethnic origin accurately, simply relying on these superficial traits. Most such characters are homogeneous within the groups and show sharp differences between the groups. This is just the opposite of what we conclude from a random, large sample of genes studied with either pre-DNA or DNA methods, where almost without exception differences between people from the various continents are quantitative, not qualitative, and of small magnitude with respect to genetic differences within groups, even in small towns or villages. For genes existing in various forms, one usually finds that the same forms exist in almost all parts of the world, but only in different percentages, and the differences in the percentages are rarely striking. Clearly, skin, eye, and hair coloration and the shape of the human face and body are not randomly distributed genes, and they show much greater racial differentiation. It seems important to discuss the reason for this difference, which is clear even if it is not widely known and which can explain this apparent contradiction. Everything we see and use in our diagnosis of ethnic origins of individuals in every day is a property of the surface of the body. Humans originally developed in the warmer parts of Earth and only later, in the last 50,000 years or so, spread to the rest of the world. They even had to learn to survive in extremely cold environments, such as Siberia. This involved a cultural as well as a biological adaptation. All the characters that show a strong difference between races, as well as greater homogeneity within races, are connected with adaptation to climate, and have been so explained by physical anthropologists. It is inevitable that they involve the surface and general shape of the body, because those aspects of people form the interface between the external world, from whose physical rigors we must protect ourselves, and the internal one, the temperature of which we try quite successfully to keep constant.

It is not surprising, therefore, that the surfaces of our bodies differ greatly depending on the climates in which our ancestors developed in the last tens of thousands of years. As the surface is the part of us that we see, its conspicuousness affects our judgment on questions of race, and prompts us to believe that all other differences are equally striking. But the truth is that they are not. We may add that they could hardly be, given that the evolutionary separation among humans living on the various continents has been relatively short; there just has not been time to develop much genetic divergence among humans.

Knowledge about our real, and very limited, genetic differences can only help to defuse the "race bomb." It is, however, obvious that any data can be misused and abused. The best protection is education based on solid facts and research.

9. The plight of indigenous populations

Many indigenous populations live in conditions of extreme poverty, many are being abused and victimized in the process of economic development. Their plight is usually well hidden, and it is very difficult to help them. The United Nations is aware of the urgent needs of many indigenous populations and has called 1993 the year of the indigenous people. Given the many terrible, urgent problems that face the world today, the limited resources of the United Nations are fully committed in

many parts of the globe. It is unrealistic to hope that this agency could do much to help indigenous populations.

The HGD Project might help indigenous populations in at least two ways. First, the Project's investigations may make the burdens borne by particular populations public knowledge. The Project can also generate greater public knowledge and interest in cultural diversity and the desirability of maintaining it. It would of course be impossible to reach all populations in need in the course of the Project; there are about 5,000 different populations in the world, based on the count of different languages in existence, and the Project can reach about ten percent of them.

10. The HGD Project and identity

The HGD Project has generated enormous spontaneous interest among journalists and laymen. The program has the potential for discovering remote origins of people, if not or not always for individuals at least for groups, and this seems to evoke strong, positive, emotional reactions. Most people know too little about their remote past, because, after their ancestors migrated to America, their family lost its collective memory of its previous country as a result of the ignorance and the anger that often accompany the poverty and oppression that forced many of them to abandon their homes. But knowledge of personal identity may be a basic interest of humans and an important component of self-esteem, and information of the kind that the program can provide may be cherished by many. It is well known that several cultures dedicate special effort to conserving their genealogies.

11. Spreading scientific knowledge

A further consideration is that the Human Genome Diversity Project can contribute to spreading interest in genetics among the public and may create a desire to learn more about our science. Greater scientific literacy should be one of our country's most important goals to prepare our population for the future. Recent international comparisons show this is a major and urgent need in the United States. The Human Genome Diversity Project will generate further knowledge of the kind promoted by the Human Genome Project, but it will also generate curiosity in many individuals, some of whom would not be interested in the Human Genome Project. This curiosity may become a powerful stimulus to additional interest in the life sciences, which are often taught in high school and will, I hope, become even more popular. My colleagues and I cannot forget—and policy makers should not forget—that science can be extremely exciting. The way the Human Genome Diversity Project connects with people interested in history, anthropology, languages, family roots, or a wide variety of topics outside molecular biology offers a wonderful opportunity for public education. The Project will give proper attention to that opportunity.

Senator AKAKA. Dr. King?

TESTIMONY OF DR. MARY-CLAIRE KING, PROFESSOR OF GENETICS AND EPIDEMIOLOGY, UNIVERSITY OF CALIFORNIA AT BERKELEY

Dr. KING. Thank you, Senator.

It is both an honor and a great pleasure to be asked to speak with you today about the Human Genome Diversity Project. As you have heard from my colleagues, the revolution in genetics offers us now the unprecedented opportunity to unlock many of the secrets of biology.

In the area of medical research, mapping and sequencing the human genome will enable us to solve the mysteries of many human diseases that have so far eluded us. As you know, my own passion in this area is breast cancer, from which women of native Hawaiian ancestry suffer the highest rates in the world.

As Dr. Collins told you, he and I and our young colleagues have every hope of identifying very soon the gene responsible for inherited breast cancer. We will know when we have found the culprit gene because the sequence of that gene will be different for women who have inherited breast cancer compared to women who have not.

How do native Hawaiian women with breast cancer differ from other women in their sequences of this gene? We don't know yet. I think we should sample native Hawaiian women and find out.

Differences in genetic sequences between people are revealing the causes of many common human diseases. All of us have the same genes. However, we often have slightly different sequences of those genes. Most of these differences have no impact on our lives, but a few do. Genes that cause terrible diseases can differ from their normal counterpart genes by very subtle differences—often only one or a few base pairs of DNA out of thousands. But the consequences of that variation can be immense. Where does this genetic variation come from? Why has it persisted? What does it mean?

Our genetic differences are at the very heart of our evolution as a species. One of the most fascinating paradoxes of the human condition is that we are all different, yet we are all the same. If I take a sample of my DNA and compare it to the DNA of another person selected at random from anywhere in the world, she and I will be the same for perhaps 999 base pairs out of 1,000, and will differ maybe at one. Because each of our cells contains 6 billion base pairs of DNA, those differences add up. There are millions of differences between two randomly selected people. But 99.9 percent of our DNA is the same.

Of course, if I select my brother and compare my DNA to his—hardly a random selection—he and I will be much more similar than two unrelated people; in fact, my brother and I will be identical for some DNA sequences.

The genetic identity of people in the same family has enabled us to bring justice to victims of human rights violations. Specifically, working with human rights groups in Argentina, my lab takes advantage of the genetic identity of specific DNA sequences among relatives to identify families of children who were kidnapped and whose parents were killed by military forces in Argentina during the Dirty War.

Working now with the U.S. Army Central Identification Laboratory in Hawaii, my lab is using exactly the same tools to identify MIAs from the Vietnam War. This approach only works because we previously analyzed the genetic diversity of the same DNA sequences. In other words, we can only establish a person's genetic identity by knowing how unlikely it would be for two unrelated people to be exactly the same for a long sequence of DNA.

At the other extreme from a brother and sister, if I select people from two populations that have been isolated from each other for most of human evolution, they will differ by more than 1 in 1,000 base pairs. But it is important to note that this variation does not follow the social concepts of race we grew up with. And understanding genetic diversity as it really exists can provide irrefutable evidence against racial stereotyping.

Dr. Cavalli was too modest to tell you, but his work demonstrated some 30 years ago that if one samples all of the residents of a small village of pygmies in central Africa, one will find in that small village virtually all of the genetic variation that one would find by taking a sample of the entire world.

Roots and origins are fascinating, but what can genetic variation tell us about the human condition today? Let me return to Africa for a medical example. Sickle cell anemia is the most common genetic disease among African Americans. We know a great deal about the genetics of sickle cell anemia, because it was possible to study this disease at the level of the hemoglobin protein and the red blood cell. Of course, most common diseases do not leave such good clues at the protein and cellular levels. That is why we need information from our genes, and that is why we need the Human Genome Project.

Until very recently, sickle cell anemia killed its victims as children or young adults. So why does it still exist? Why didn't the sickle cell gene die out with its victims? As any young biology student will tell us, sickle cell anemia continues because people who carry the sickle cell gene are protected against malaria. It only takes one copy of this gene to protect a child against malaria, but it takes two copies for a child to develop sickle cell anemia.

Malaria has killed more people in human history than any other disease. So when a DNA sequence appeared in human evolution that protected children against the most powerful killer in their midst, that gene flourished, even though two children protected by that gene might grow up and have a child together who inherited a sickle cell gene from each of them and died of sickle cell anemia.

Now, how was this complex story of hemoglobin, sickle cell anemia, and malaria disentangled 50 years ago? It was not done by studying modern African Americans. Malaria no longer exists here. Rather, the puzzle was solved by studying indigenous populations from areas of Africa where malaria had been common since the beginning of agriculture and where people had lived in relative isolation since that time.

The correspondence between malaria and sickle cell anemia among populations in central Africa was very strong. Populations with high rates of malaria also had high rates of sickle cell anemia. We now know at the genetic level that people with sickle cell anemia have a specific mutation in a hemoglobin gene.

The point of this story is that sickle cell anemia was understood by studying the ancestral African populations from which modern African Americans are descended and by analyzing the genetic variation among those African populations.

I would tell you the story of blindness on the island of, but I realized this morning that you know more about it than I do, so I will go on.

The Human Genome Diversity Project is still evolving. We still have a great deal of work to do. In the past year, our evolution has been quite rapid as a project. We have developed an administrative and scientific infrastructure for the project and have held three workshops. You have heard about these workshops from Dr. Marrett. The first was devoted to mathematics; the second to anthropology, which populations best represent our common human ancestry, which are in most danger of extinction, what are the linguistic relationships among these populations?

Very importantly, in this country, our indigenous ancestral populations are not only subjects of the study, but also coworkers, whose participation is crucial for our success. Their cultural and

linguistic resources will allow us, in fact, to work successfully on Native Americans in the project. Representatives of Native American tribes have contacted us and asked to participate collectively in the work.

The third workshop at Bethesda was the longest and was devoted to ethical, legal and social implications of the project. How are the rights of people who blood samples to be protected? What is informed consent under these circumstances? How can we provide open access to materials for researchers and at the same time protect against misuse?

These same issues faced us and continue to face us as researches in the larger Human Genome Project. We look forward to working with our colleagues in the Human Genome Project in trying to understand and address these very thorny subjects.

I would particularly call your attention to the report of the workshop on ethical, legal and social issues, prepared by Professor Hank Greely from the Stanford University Law School. It is included with your materials. That workshop established a committee structure consisting of four subcommittees—on collection, social implications, review and evaluation for the Human Genome Project. Of particular interest was the concept of a set of rapid-response committees that could address possible misuse of genetics, both from the Diversity Project and generally, should such allegations arise.

The second issue we addressed at the Bethesda workshop was how to select genetic markers which will provide the most scientific information for the questions we want to answer. This addresses some of the concerns that were raised by my scientific colleagues a few minutes ago.

The third issue at the Bethesda workshop was the question of how the Diversity Project complements the research interests and programs of the National Science Foundation, the National Institute of General Medical Sciences represented here this afternoon by Ruth Kirschstein, and the Human Genome Project.

We enthusiastically support a review and discussion of the Human Genome Diversity Project by the National Research Council and look forward to working with the Human Genome Project personnel on ethical, legal and social issues shared by both projects.

In closing, let me pose a few of the fascinating and important questions of human experience that can be addressed by studying human genome diversity.

When people have the same exposures to viruses or bacteria, why do some get sick and others not?

Are high blood pressure and diabetes common now in some American populations because genes for these conditions were an advantage in ancestral populations earlier in human evolution?

How do people migrate? How did humans move out of Africa? Were the Americas settled in waves or in streams? How was Hawaii settled? By whom? How do languages and cultures move with people?

In summary, the genetic diversity of living humans reflects the evolution of our species. We believe we can link molecular analysis of the human genome with population genetics, anthropology, archaeology and linguistics to study our past and present. The goal

of the Human Genome Diversity Project is to understand who we are as a species and how we came to be.

Thank you. I'd be delighted to answer questions.

PREPARED STATEMENT OF DR. MARY-CLAIRE KING

It is both an honor and a great pleasure to be asked to speak with you today about the Human Genome Diversity project. As you have heard from my colleagues, the revolution in genetics offers us now the unprecedented opportunity to unlock many of the secrets of biology. In the area of medical research, mapping and sequencing the human genome will enable us solve the mysteries of many human diseases that have so far eluded us. My own passion in this area is breast cancer, from which, as you know, women of native Hawaiian ancestry suffer the highest rates in the world. As Dr. Collins told you, he and I and our young colleagues have every hope of identifying, very soon, the gene responsible for inherited breast cancer. We will know when we have found the culprit gene, because the sequence of that gene will be different for women who have inherited breast cancer compared to women who have not.

Differences in genetic sequences between people are revealing the causes of many common human diseases. All of us have the same genes. However, we often have slightly different sequences of those genes. Most of these differences have no impact on our lives. But a few do. Genes that cause terrible diseases can differ from their normal counterpart genes by very subtle differences—often only one or a few base pairs of DNA out of thousands. But the consequences of that variation can be immense. Where does this genetic variation come from? Why has it persisted? What does it mean?

Our genetic differences are at the very heart of our evolution as a species. One of the most fascinating paradoxes of the human condition is that *we are all different, yet we are all the same.* If I take a sample of my DNA and compare it to the DNA of another person selected at random from anywhere in the world, she and I will be the same for about 998 base pairs out of 1000 and will differ at only two. Because each of our cells contains 6 billion base pairs of DNA, those differences add up to millions between two randomly selected people. But 99.8 percent of our DNA is the same. Of course, if I select my brother and compare my DNA to his (hardly a random selection), he and I will be much more similar than two unrelated people. In fact, my brother and 1 will be identical for some DNA sequences.

The genetic identity of people in the same family has enabled us to bring justice to victims of human rights violations. Specifically, working with human rights groups in Argentina, my lab takes advantage of the genetic identity of specific DNA sequences among relatives to identify families of children who were kidnapped and whose parents were killed by military forces in that country during the Dirty War. Working with the US Army Central Identification Laboratory in Hawaii, my lab is now using exactly the same tools to identify MIAs from the Vietnam War. This approach only works because we previously analyzed the genetic diversity of the same DNA sequences. In other words, we can only establish a person's genetic identity by knowing how remarkable it would be for two unrelated people to be exactly the same for a long sequence of DNA.

At the other extreme from a brother and sister, if I select people from two populations that have been isolated from each other for most of human evolution, they will differ by more than 2 in 1000 base pairs. It is important to note that this variation does not follow the social concepts of race we grew up with, and understanding genetic diversity as it really exists can provide irrefutable evidence against racial stereotyping. Let me explain. Some very interesting work has been done with DNA sequences from people of African ancestry, people of European ancestry, and people of Native American ancestry. DNA sequences from two randomly selected African individuals will differ from each other *more* than an African's sequence will differ from a Caucasian European's sequence or a Native American's sequence. That is, there is more genetic variation within Africa alone than there is between Africa and anywhere else. Why? Because human evolution has its most profound depth in Africa. As a species, we have lived longest there. Fundamentally, we are all African. Understanding our genetic roots reveals our common human origin.

Roots and origins are fascinating, but what can genetic variation tell us about the human condition today? Let me return to Africa for another medical example. Sickle cell anemia is the most common genetic disease among African Americans today. We know a great deal about the genetics of sickle cell anemia, because it was possible to study this disease at the level of the hemoglobin protein and the red blood

cell. (Most common diseases do not leave such good clues at the protein and cellular levels. That is why we need information from our genes.)

Until very recently, sickle cell anemia killed its victims as children or young adults. So why does it still exist? Why didn't the sickle gene die out with its victims? As any young biology student will tell us, sickle cell anemia continues because people who carry the sickle cell gene are protected against malaria. It takes only one copy of the sickle cell gene to protect a child against malaria, but it takes two copies of the sickle cell gene for a child to develop sickle cell anemia. Malaria has killed more people in human history than any other disease. So when a DNA sequence appeared in human evolution that protected children with one copy of it against the most powerful killer in their midst, that gene flourished, even though two children protected by that gene might grow up and have a child together who inherited a sickle cell gene from each of them and died of sickle cell anemia.

How was this complex story of hemoglobin, sickle cell anemia, and malaria disentangled 50 years ago? It was not done by studying modern African Americans. Malaria no longer exists here. Instead, the puzzle was solved by studying indigenous populations from areas of Africa where malaria had been common since the beginning of agriculture and where people had lived in relative isolation since that time. The correspondence between malaria and sickle cell anemia among populations in central Africa was very strong. Populations with high rates of malaria also had high rates of sickle cell anemia. We now know at the genetic level that the people with sickle cell anemia have a specific mutation in a hemoglobin gene. The point of this story is that sickle cell anemia was understood by studying the ancestral African populations from which African Americans were descended and by analyzing the genetic variation among those African populations. Only among these ancestral populations who lived and still live in geographic areas where malaria was so devastating was the sickle cell gene an advantage, even though in a double copy, it was lethal. The ancestral populations provided the critical clues, and proteins were the critical molecules. Ancestral populations still provide critical clues, with molecular evidence now coming from our genes.

The Human Genome Diversity project is still evolving. We still have a great deal of work to do. In the past year, our evolution has been quite rapid. We have developed an administrative and scientific infrastructure for the project and have held three workshops. The first workshop, at Stanford, California, was devoted to mathematics. How many samples need to be gathered and how many DNA sequences or markers need to be typed for each person? The second, at University Park, Pennsylvania, was devoted to anthropology. Which populations best represent our common human ancestry? Which are in most danger of extinction? What are the linguistic relationships among these populations?

The third workshop, at Bethesda, was the longest and was devoted to three issues. First were the ethical, legal, and social implications of the project. How are the rights of the people from whom a blood sample is requested to be protected? What is informed consent under these circumstances? How can we provide open access to materials for researchers and at the same time protect against misuse? I would particularly call your attention to the report of that workshop prepared by Professor Hank Greely from the Stanford University Law School, which is included with your materials. Second was the question of selecting the genetic markers which will provide the most scientific information for the questions we want to answer. The third issue at the Bethesda workshop was the question of how the Human Genome Diversity project complements the research interests and programs of the National Science Foundation, the National Institute of General Medical Sciences, and the Human Genome Project. We have been working with these agencies and believe the Human Genome Diversity project reflects the commitment to biological anthropology at the NSF, contributes to the development of resources that has been so historically important at the NIGMS, and can both contribute to and take advantage of technological developments in rapid typing and sequencing and in data base management at the Human Genome Project.

In closing, let me pose a few of the fascinating and important questions of human experience that can be addressed by studying human genome diversity. When people have the same exposures to viruses or bacteria, why do some get sick and others not? Are high blood pressure and diabetes common now in some American populations because genes for these conditions were an advantage in ancestral populations earlier in human evolution? How do people migrate? How did humans move out of Africa? Were the Americas settled in waves or streams? How was Hawaii settled? By whom? How do languages and cultures move with people?

In summary, the genetic diversity of living humans reflects the evolution of our species. We can link molecular analysis of the human genome with population genetics, anthropology, archaeology, and linguistics to study our past and present. The

goal of the Human Genome Diversity Project is to understand who we are as a species and how we came to be.

Senator AKAKA. Thank you very much, Dr. King. Both of your testimonies will be helpful to us.

I will address the following questions to the panel as a group. The agency witnesses have raised some very interesting points that I believe need to be addressed. I will prepare some written questions on this topic which I will ask you to answer for the record.

Representatives from various Federal agencies have highlighted their position on the importance of the Diversity Project. This is an opportunity for you to demonstrate that the Diversity Project represents a commitment to advancing medical science. My question is: How might the Diversity Project help medical researchers understand why different populations are susceptible to different diseases?

Dr. KING. It is well-known among epidemiologists and medical researchers that in some populations, a particular disease will be very common, whereas the same disease in other populations will be very rare.

For example, in parts of Finland which have been genetically isolated for a long while, myocardial infarction is very common. If we can understand why that is true at the level of our genes, that information will inform us not only about heart disease in Finland, but also about heart disease for all of us. The biology of the disease is the same. But we have the particularly good fortune to be able to pinpoint populations in which genetic influences on the disease both occur and are frequent.

Isolated populations are particularly valuable for this purpose. As Dr. Collins mentioned, one needs a larger sample than just 25 individuals from each population in order to carry out such an analysis fully. That is why we propose to select a few individuals from each population from whom we would like to make permanent access of DNA available; but many more individuals from each population so that we can close in on exactly these specific questions as they arise.

Senator AKAKA. Can you cite examples of how the study of ancestral populations have increased our understanding of health problems afflicting modern Americans? Do you suggest that the study of the genetics of ancestral populations might shed light on medical issues that are relevant today?

Dr. KING. Sickle cell anemia is an interesting example of how the investigation of a disease in ancestral populations can tell us a great deal about why that disease is so common in America today.

Another example might be the investigation of the autoimmune disease systemic lupus erythematosus, which is very common among Americans of African ancestry and among Americans of Hawaiian and other Pacific ancestries, but is less common among Americans of European ancestry. If we can investigate the genes responsible for susceptibility to that disease in the Pacific and in Africa, we will learn a great deal about the underpinnings of that disease generally and will have much better tools for diagnosis and treatment of it.

Senator AKAKA. Very interesting. The Human Genome Project under the sponsorship of NIH and DOE has tried to establish a

single reference genome. This reference genome will be made up of many individual DNA samples, but would not contain the level of diversity that you advocate. What medical science opportunities might be overlooked if the Human Genome Diversity Project does not proceed?

Dr. CAVALLI-SFORZA. This is where we still don't know enough to give a totally satisfactory answer, but we can give examples that already show that it is going to be very important. What little we know about human variation from one individual to another shows that it is very different, depending on the particular part of the genome being analyzed. There are parts of the human genome that are almost invariable from one individual to another. They usually are important in the sense that if they vary, the individual may be dead or severely sick. So that knowledge is important from a medical point of view.

But there are also other parts of the genome that are found to vary some time considerably from one individual to another. We know that some vary without detriment, but others must vary for individuals to function normally. And recently variation was detected which might have seemed trivial, but it turns out that it may be responsible for important diseases, one of which is Huntington's Chorea.

Senator AKAKA. Preservation of genetic data of near-extinct populations is an important issue that I believe warrants further examination. Would you say that a primary component of your proposed study is the collection of genetic samples from cultures that are vanishing? Why does this project need to be done now? Why the urgency?

Dr. CAVALLI-SFORZA. We are observing, because of economic development of Third World countries, an enormous increase in migration, and also disappearance of populations living in peripheral locations, and very frequently in marginal conditions. The increase in communication and the opportunities of working abroad are creating enormous changes in human customs and peoples distribution and admixture.

If we want to reconstruct populations history, and understand under what conditions certain diseases became prevalent in a population, we have to study these things now, before complete confusion. So I believe that we really have to start being active very soon.

Senator AKAKA. The Human Genome Project currently funds research that investigates the legal, ethical and social implications of their ongoing genetic studies. What measures do you feel must be taken to ensure that a diversity study is sensitive to the concerns of the various ethnic groups about privacy and racial stereotyping?

Dr. KING. There are three classes of issues here, all difficult. I might call those issues rights, royalties and racism.

In the area of rights, the question is: What constitutes informed consent in isolated populations? In this area, we confront the same concerns that confront us in working in the larger Human Genome Project. We will establish a procedure for model protocols for working with populations from widely diverse cultures so that standards that are coherent with the culture are employed for the projects we hope to undertake.

It is absolutely critical to us that representatives of the cultures that we hope to investigate be part of this process; that people not simply be the subjects of studies, but participants in them.

In the area of royalties, we need expert advise. Dr. Greely is in the process of working with legal experts in ownership of information questions and how those apply to the Third World.

On the question of racism, I must tell you that I entered human genetics because the question of race and IQ reared its ugly head in Berkeley when I was a graduate student there, happily working in theoretical statistics. I realized that it was simply not possible to be a scientist in a vacuum.

That specific question may plaque us less now, but the overall question of racism and the misuse of genetic information as racist propaganda is certainly still with us, and nowhere is that more obvious just now than in Europe.

We need to balance in research generally and in this project in particular the need for openness of information and open access to resources without censorship against the danger of misuse.

I am not sure any of us will ever have the perfect solution. An open society leaves open the possibility of demagoguery. One approach that we discussed at the workshop is to have ready response committees available, people who can be called upon in genetics, in social sciences, in law and in ethics, at a moment's notice, to address outrageous claims if they arise.

We do not wish to shut off access to information, but we most certainly wish to put debate about these concerns into their proper scientific perspective.

Senator AKAKA. Does the Human Genome Diversity Project have a position on Federal funding and support for the Human Genome Project?

Dr. KING. The Human Genome Project is proceeding remarkably rapidly, despite having been underfunded since its inception. It is going more quickly than anticipated based on full funding, and it has done so on far less than full funding.

It is essential that the core goals of the Human Genome Project proceed as originally scheduled.

What we hope to do is integrate the goals of this project with the goals of the Human Genome Project, particularly in the area of technology development, and in addition to obtain funds that will enable both sets of goals to be carried out more rapidly. From our point of view, it is essential that both the larger Human Genome Project and the far smaller Human Diversity Project be more fully funded so that we can get answers to these questions while it is still possible to do so.

Senator AKAKA. The President has launched his technology initiative to help the United States regain our competitive edge in many high-tech industries. A key component of the initiative is education. How will the Human Genome Diversity Project contribute to public education about genetics?

Dr. CAVALLI-SFORZA. As I already indicated briefly in my presentation, I believe that the Human Genome Diversity Project is a perfect example for education in genetics, because people are very interested in their origins and history. The Project offers a natural way to bring them to want to study the subject of genetics.

There already exists one example that has been very successful. The Musee de l'Homme has started an exhibition in which it explains human individual diversity and racial diversity. This exhibition has had enormous success and has travelled to many other parts of Europe where, as I am sure you know, there have been explosions of racism which are becoming very worrisome to many of the governments.

I might add perhaps a few words on what you were asking before about racial stereotyping. The real diversity we find is between individuals within a group, but not so much between groups. We are misled by an unfortunate fact, in that we see just the surface of people, the surface being skin color, and the size, and the general body shape and facial shape, etc. But that has been designed by natural selection to respond to differences in climate, and it does not respond to any other thing, except it is the only thing we see. So from the fact that blacks are black, whites are white, et cetera, more or less uniformly, we argue that every other trait is equally uniform between races, but this is absolutely wrong, and this is the message we already know very well and should be made much more widely known. We also know it is not going to be contradicted by the DNA data, because we already have a substantial amount of information from DNA confirming what we already know from pre-DNA studies. If we make this known on a very wide scale, we will really help to eliminate at least some of the prejudices that accompany and cause racism.

Senator AKAKA. I liked the remarks made by Dr. King and yourself about educating people like me. To find something that will unlock the secrets of genetics is fascinating, and something that defines the evolution of our species is very interesting. It is an amazing project that you are engaged in.

Finally, as scientists, what do you feel is the most important discovery that we can learn from the Human Genome Diversity Project?

Dr. CAVALLI-SFORZA. Maybe summarizing what I was saying, we can learn that we individually are extremely diverse; as groups, we are not as diverse as we think we are; we are much more similar, and our kinship is extraordinarily high. But also from that individual diversity, we may learn something useful for an orientation toward our best behavior in terms of our nutrition, in terms of our protection against diseases, and this will require learning more about our individual diversity.

Senator AKAKA. Do you have any remarks on that, Dr. King?

Dr. KING. As always, Luca is a very hard act to follow. I think the understanding of how we are all the same and yet, all different, will enable us to understand a great deal about why we are differentially susceptible to different diseases. In particular we can learn how genetic susceptibility and environmental exposures interact with one another for many of the most acute viral diseases that we confront today.

Senator AKAKA. I want to thank you and the other witnesses for participating in today's hearing. I would rate the hearing a success, and I say that because I believe we have learned a great deal—I have—about the scientific merits of the Human Genome Diversity Project. I have a better appreciation for the Diversity Project and

feel that the Federal Government should strongly consider supporting this project. I hope that my colleagues will feel the same when they read the record and learn more about what you have said.

I am confident that concerns about the misuse of genetic data collected during the study will remain a central concern to the Diversity Project team members. I stress that the effort to avert racial bias and stigmatization must remain a primary concern for both projects.

I hope that the agencies present for today's hearing will seriously consider the benefits that the Diversity Project offers. The Diversity Project is motivated by the urgency, as expressed here, to collect the genetic samples before it is too late.

I will review the information gathered today and will closely follow developments related to the Human Genome Project and the supplemental Human Genome Diversity Project. Again, I thank all the witnesses for your excellent testimonies and recommendations. The Committee will keep the record open for ten working days to receive additional written testimony.

If there are no further questions or comments, then this concludes today's hearing. The Committee will stand in recess subject to the call of the chair.

[Whereupon, at 3:42 p.m., the Committee was adjourned.]

APPENDIX

THE PENNSYLVANIA STATE UNIVERSITY
University Park, PA, April 30, 1993.

Hon. Daniel Akaka, Senator
Washington, DC

Attn: Shane Merz

DEAR SENATOR AKAKA: Everyone knows the poignancy of extinction. Natural history museums around the world exhibit specimens of extinct beings. The only remnant we have of many species are a few specimens collected in the past few centuries by dedicated naturalists, and we are grateful to them for preserving those specimens, and to our museums for making them available to us—and to posterity. It is all we have left to see, to remember.

Something similar applies to our own species. Human populations around the world are systematically being absorbed into the greater mass of large national populations, their ethnic and biological identities disappearing forever. One aim of the Human Genome Diversity project is to collect and preserve for posterity adequate samples of the unique genetic variation found in and between the diverse populations of our species.

Such data can help anthropologists, linguists, archeologists, and human biologists to reconstruct the history and pre-history, settlement origins, and ethnic patterns of different regions in the world, and to understand the kaleidoscope of our similarities and differences. These are the anthropological objectives of the Human Genome Diversity Project.

If our country has entirely lost its will to do basic science—things done for interest and for their own sake—we will lose our leading place in current world science, and our noble place in world *history* of science, and will rightly be blamed by generations of our children for thinking only of our own immediate, selfish, material interests.

However, if the effort to preserve our species' genetic patrimony, with all of its diversity, is important for its own sake, there are also immediate practical benefits to the HGD project.

Human genetics is making incredible strides in our understanding of the genetic factors that are involved, to a greater or lesser extent, in nearly every human disease. We owe much to the efforts associated with the Human Genome Project. But if there is one thing we are learning from that Project, it is that even for the simplest diseases there is a large amount of variation within any human population: cases of the same disease in different individuals are caused by different genetic variants and, importantly, different genetic variants produce clinical differences in the disease, and these can be important in treatment, prevention, and prognosis.

Within a given population, a few variants are involved in most of the cases, but a substantial number of other rarer variants are also found. Where we have adequate data, even closely related populations (e.g., different European nations) can have largely nonoverlapping sets of causal variants for a given disease. Only after we have studied a population thoroughly, and know its specific variants, can we work on the treatment approaches relevant to each, and do effective risk counseling.

Moreover, different ethnic groups have *different genetic variants* for the same disease, and sometimes these do not overlap at all. We have a lot of data on Europeans, Japanese, and so on, that is, about people living in wealthy places with major laboratories. But how do we understand the risk of disease in other groups? Even if we know the gene and its function, we must search each ethnic group separately to identify that group's disease-causing variants.

This country has among its citizens numerous people from all but the smallest and most isolated human populations. Our big cities like New York and Los Ange-

les, and our states, like Alaska, Arizona, and Hawaii, are inhabited by very diverse populations, and these include large numbers of people with admixed ancestry among different populations. One need think only of Mexican-Americans, with European-Amerindian ancestry, or African-Americans, to remember that there are *tens of millions* of these people in our population. In aggregate, they may soon comprise the majority of our citizens. For various diseases, these people are subject to risks that differ, probably at least in part for genetic reasons, from the risk in the better-studied part of our nation.

The National Institutes of Health *mandates affirmative action* in all of its research grants, to ensure that all of our nation's people are served by the research and clinical establishment. How is it that we can allocate vast sums of money to studying human genes in a way that specifically *excludes* consideration of relevant diversity?

The organization (and a representative sequence) of our genome can be documented any time and without urgency. We will benefit greatly from this being done. But many peoples in the world are facing inevitable loss of their genetic integrity as populations. A worldwide sampling of human genetic diversity, including rare and endangered populations as well as the major ethnic groups, would cost only a trivial fraction of the amount being spent to understand the organization of the genome, would preserve this variation for posterity.

The core of the proposed Genome Diversity Project can document this diversity in general terms. A subset of larger samples, or even disease-related screens, focused on those populations that comprise our major ethnic groups (proper representation of Africans, Amerindians, Pacific and Asian populations, Indians, etc.), can provide biomedically critical data.

These biomedical objectives are important aspects of the Human Genome Diversity idea. But I think it is just as important that the project, with its inherent, social, historical, and ethnic interest, can help preserve for American science a semblance of scientific inquiry that is rapidly being lost amid the ever more blatant scramble for funding to do immediately 'relevant' work.

This letter itself represents part of the sad need we face in today's science to devote large fractions of our time to seeking funds, and to try to justify the expenditure. I think our country should again make room for scientific inquiry that is justified because it is interesting—that is what drew most of us to scientific careers in the first place. It is the basis of our scientific research *and training* infrastructure, and the seed-bed for new ideas. It is willingness to nurture free curiosity that made this country the scientific envy of the world.

If I can provide you with more specific information and ideas about the Genome Diversity initiative, please let me know.

Yours truly,

KENNETH M. WEISS, Ph.D.
Professor of Genetics & Anthropology
Head, Department of Anthropology

ANSWERS TO QUESTIONS FROM SENATOR AKAKA

After hearings of the Senate Committee on Government Affairs concerning the Human Genome Diversity Project (the "HGD Project"), held on April 26, 1993, Senator Daniel K. Akaka asked Professor Henry T. Greely on behalf of the HGD Project Organizing Committee and particularly its two witnesses before the Committee, Professor L. Luca Cavalli-Sforza and Professor Mary-Claire King, to respond for the record to several written questions. These materials provide the answers sought, as well as commenting on a few other matters of note concerning the HGD Project. The HGD Project Organizing Committee is grateful for this opportunity, and the opportunity afforded by the hearings themselves, to expand Congressional and public knowledge of this exciting scientific endeavor. We would be happy to answer any further questions concerning the HGD Project.

Question 0.1. Please outline the immediate goals and objectives of the Human Genome Diversity Project.

Answer. The goal of the Human Genome Diversity (HGD) Project is to help us better understand our species through collecting, preserving, and analyzing genetic material from diverse populations around the world.

The investigation of individual genetic variation in the groups studied will have the following objectives:

a) to help reconstruct the history of the world's populations, including important demographic events such as major migrations and expansions;

b) to obtain information of potential or actual medical interest for epidemiological objectives, directed to the health and welfare not only of the individuals or groups sampled but also of the rest of the world;

c) to explore, as far as is practical, biological problems such as, for example, the functions of genes or the adaptations of humans to different environments, as well as to provide a resource for future work on these points; and

d) to provide a basis for comparisons with similar data on DNA from archaeological material.

To achieve these aims, the Project will undertake the collection of biological samples from a number of human indigenous populations, paying special attention to endangered groups but also including appropriately sampled individuals from less remote areas, in order to obtain a balanced picture of human genetic variation. At the very beginning of the Project, some of the samples may come from indigenous populations that are relatively easy to reach (for example, in North America), but the most endangered populations will be considered as early as possible.

Second, the Project needs to begin the transformation and preservation of the samples. White blood cells collected from an adequate number of individuals will be transformed and stored using a technique that can provide essentially unlimited amounts of DNA, so that they can immediately be used for research and also preserved for future studies. Without transformation, the value of the collections would be greatly limited. We plan to store transformed cell lines at the Coriell Institute for Medical Research in Camden, the existing cell-line facility established by the NIH Institute for General Medical Sciences (NIH–GMS).

DNA thus collected will be made available for analysis to interested scientists, and every attempt will be made to ensure that the use of this material will be as efficient and as scientifically valuable as possible. Results of DNA analysis will be stored in appropriate data bases available to scientists.

The HGD Project is in the process of completing its organizational structure and planning. This includes forming committees to oversee the collection, storage, and analysis of samples. It will also ensure the stimulation of similar initiatives in other countries and parts of the world and coordination and cooperation with them. There is a European HGD Project that has already received initial funding from the European Economic Community. Another committee structure will address the ethical, social, and legal issues in depth.

Three workshops held in the past ten months in preparation of the HGD Project have been financed by the National Science Foundation, the National Institutes of Health, and the Department of Energy. While the planning for the science and organizational structure is proceeding, the Project is also searching for funding.

Question 1.1. What criteria should be employed to create a final sampling list to ensure that the Human Genome Diversity Project is consistently implemented?

Answer. The HGD Project will give highest priority to indigenous populations that are more likely to disappear because of actual extinction or absorption into other groups and to isolates of historical interest. Frequently, these two criteria apply to the same population. The collection should not be limited exclusively to these groups, and the survey should eventually represent all the world's indigenous populations as fully as possible from geographic, ethnic, and linguistic points of view. In collaboration with 55 anthropologists at the second planning workshop, we have generated a list of about 500 populations of special interest. These represent roughly one-tenth of those that can be recognized as linguistically distinct. This list is tentative and may be adjusted as the work proceeds.

We should start casting our net on the whole world in a systematic way by spacing locations to be sampled so that, even if they are not exactly equidistant from one another, the most important regions and people will be sampled first. In this way, in a short time we can develop an approximately representative world survey, which can be sequentially improved. The initial, sparse network of locations will thus be filled progressively, making use of the information and experience as it accumulates from the survey itself. Some priority will be given to populations in the United States, both to test the collection routine under less difficult conditions and to reassure other countries that American researchers have done the same research with American indigenous populations.

Question 1.2. Besides collecting blood samples, what other biological material should be collected for DNA analysis? Should this material be collected from just the 25 individuals from whom blood was drawn? Since useful information can be derived from DNA samples alone, how many additional human research subjects are proposed for supplemental sampling?

Answer. Transforming white blood cells in order to obtain indefinite amounts of DNA is an expensive procedure. Our plan, based on discussions and consensus from the first workshop, is usually to limit transformation to about 25 individuals per population. Non-transformed genetic material is much less expensive to store. The Project will obtain that kind of material—blood samples for direct storage, hair roots, and cells dislodged by gentle scraping inside the subject's mouth—not only from the 25 individuals whose transformed cells will be stored, but also from 100 or more subjects in the sampled population or its immediate neighbors.

Methods of amplifying the limited amounts of DNA thus obtained (that is, augmenting the number of DNA copies) will also be developed and tested, but with our present knowledge, transformed cell lines offer the most satisfactory method of preservation.

Question 1.3. What is the best method for storing and assuring the continued quality of these blood samples?

Answer. The most widespread and reliable method of transformation of white blood cells for the purpose of generating indefinite amounts of DNA is the treatment of B lymphocytes (a special type of white blood cells) with Epstein-Barr virus. Current investigations of genetic diseases use this procedure, and a repository of transformed cells (where we plan to house our collection) has been established for this purpose at Camden, New Jersey, by NIH–GMS. After preliminary cell multiplication, B lymphocytes are stored in liquid nitrogen, and samples are taken when necessary for further multiplications. A very small fraction of DNA, involved in making antibodies, apparently is altered in these cells. For this reason it may also be important to store, without transformation, other cells from blood or tissues in which this part of the DNA is unchanged.

Question 1.4. Immortal cell lines will require perpetual upkeep. What would be the long-term costs associated with this Project?

Answer. A very large number of cell lines can be stored in relatively limited space, and it is inexpensive to replace the evaporating liquid nitrogen used to keep them at the right temperature. According to the Coriell Institute for Medical Research in Camden, after the initial project period, the costs for perpetual storage at May 1993 prices would be about $4 per individual per year (including overhead), based on the amount of material we now think should be stored per individual. To maintain the full 10,000 cell lines that the Project contemplates—at least 25 individuals from each of 400 populations—would cost about $40,000 per year. This is equivalent to the amount that could be provided indefinitely by an endowment of $1 million, earning a 4 percent return. That million dollars is only a few percent of the 5 year cost of the HGD Project as we see it today.

Question 1.5. What kind of data bases will be necessary to manage the storage and retrieval of biological material?

Answer. A data base will be needed to store the information on the origin of the various groups tested. This will include references to the relevant ethnographic and other anthropological information already available about the groups examined or gathered by the scientists involved in the expeditions.

Another, much larger data base will contain all the data on DNA segments tested from the individuals of the survey. The need for such a data base is being considered by the HGD Project, but any such data base should be compatible with the Genome Data Base of the Human Genome Project, funded by the National Institutes of Health and the Department of Energy.

Ideally, the same base pairs (the elementary units of DNA) should be examined on all the transformed samples from all the individuals, a highly desirable task but difficult to realize in its entirety. The size of the data base on DNA variation may be large, since it will be determined by the number of individuals tested times the number of base pairs tested on average per individual. The actual number of base pairs examined per individual may be considerably increased by technological developments likely to take place in the next two or three years thanks to the Human Genome Project. The new technologies are almost certainly applicable very directly and in a highly efficient way to the HGD Project. It is possible that when the collection of population samples has been completed, the total amount of information in the data base on human diversity might be of a size not far from that of the human genome itself, about 3 billion base pairs.

Question 1.6. What minimum set of genetic markers should be analyzed for all samples? In what priority will samples be analyzed once a common set of markers will be established?

Answer. We don't have a precise answer yet, but in general, for reconstructing history, it is best to evaluate several hundred variable genetic sites. The experience accumulated so far shows that evaluating a large number of sites greatly enriches the analyses that are possible and consequently our understanding of human diversity. The great advantage of cell lines is that one can always return to them for more information.

The analysis of variation can be carried out for single base pairs (the units forming DNA), or for whole segments of DNA made of hundreds or thousands of base pairs. The latter approach is especially productive for those segments of DNA that are particularly variable. We already know some of these, and know some rules about where and why to expect more or less genomic variation, but more must be learned in order to accomplish a rational selection of the most appropriate DNA segments. These segments should be a representative sample of the whole genome.

Individuals and markers will be analyzed with priorities dictated by the amount of information they can provide.

Question 1.7. Will 25 samples from each population be sufficient to gain important information about medical genetics?

Answer. At the second workshop, we decided that collection of samples from 25 individuals per population was the best compromise between the need for more populations and the desire to have more individuals per population Although this is adequate for most questions currently anticipated, it is not sufficient for investigating all questions, especially those that focus on within-population variation at a specific gene. This issue is being addressed in four ways: (1) the collection and storage of DNA of additional individuals from most or all of the populations, (2) the use of independently existing collections of stored cells and DNA, whether transformed or not, which may be donated to the HGD Project, (3) research work on establishing libraries of amplified DNA from non-renewable samples; and (4) establishing larger numbers of cell lines from a selected subset of populations. Items one and three were addressed briefly in the answer to Question 1.2. These additional resources will allow selected studies, and especially those of greatest medical relevance, to be carried out on a much larger sample from a number of populations.

Question 2.1. Should any benefits of the research accrue to the research subjects?

Answer. The donors of the blood and other DNA samples will benefit from their participation in the HGD Project in several ways. First, the Project may lead to improved information about medical problems afflicting the donor populations. Second, the Project would like to require that the developers of any commercial products created from the Project's samples pay royalties to a fund to be used to benefit the participating populations. Third, the Project intends to transfer biological technology to other parts of the world, thus providing some scientific and economic development in some of the countries from which samples are taken. Fourth, for many small and threatened populations, participation in the HGD Project should increase the general public knowledge of their existence and the dangers they face. This improved visibility may indirectly help protect them from a variety of outside pressures.

Finally, in many cases the Project may be able to offer more immediate and tangible benefits to the populations through the provision of some immediate medical care. In Dr. Cavalli-Sforza's work with African pygmy populations, for example, he learned that 3 to 4 percent of the residents in camps he visited were afflicted with yaws, a destructive skin disease. He arranged, with assistance from drug firms, to bring them penicillin, which treats that disease very effectively. The kind of medical care that could be offered will necessarily vary with the circumstances, but the Project will encourage that kind of humanitarian return to the populations that are sampled.

Question 2.2. What are the risks, particularly social risks such stigmatization, to research subjects who participate or decline to participate?

Answer. The risks will necessarily vary from culture to culture, which is one of many reasons that the Project intends to sample only populations for which anthropologists or others have established a good understanding of the culture. Working with the experts in the culture, we will strive to arrange the donations so that no one would be harmed by his or her decision about participating. In fact, though, this kind of research has been done many times, and the researchers involved believe that this particular threat is extremely unlikely. Drawing blood is not a novel procedure for research in isolated populations. In almost all such populations, some people are willing to donate blood and others are not. Neither decision would normally put a person into a tiny and threatened minority.

Question 2.3. What limits are necessary to assure the confidentiality of the samples collected?

Answer. The identities of the individuals from whom samples are collected will not be publicly available, nor will the samples stored at the repository be identifiable by individual names. We foresee no difficulty in protecting the subject's privacy or in following all applicable Federal law on confidentiality.

Question 2.4. The issue of informed consent raises its own questions: What constitutes meaningful informed consent in non-Western cultures? Do notions of informed consent have any true relevance to some of the populations sampled?

Question 2.5. How will the Human Genome Diversity Project comply with the "Federal Policy for the Protection of Human Subjects" (10 CFR Part 745)? What difficulties might the project encounter subscribing to the human subject protection regulation?

Answer. We believe it is most useful to answer these two questions together. There is a fascinating debate among researchers and ethicists about the meaning and propriety of "informed consent" in some cultures. Some have argued against the "ethical imperialism" of subjecting other cultures to American norms of disclosure and consent; others have urged that these are issues of universal human rights. The HGD Project does not need to take a position on this debate. Federal law requires informed consent, in fairly sharply defined terms; medical-anthropological research on isolated populations has proceeded successfully in the past subject to these legal requirements; and the HGD Project will comply fully with Federal law as well as the laws of the host countries.

We do believe that the idea of informed consent is relevant for all human populations. It embodies a respect for the individual that should be applied universally. The method of seeking informed consent necessarily will vary from culture to culture, within the constraints of Federal law. How to describe the Project and its research in a way that has meaning to members of an individual population will vary. This again points up the vital role anthropologists and other experts on local populations must play in the Project.

The Federal regulations contain details that will be of limited value in some settings. For example, they might require the Project to provide written documents to populations that are entirely illiterate. We do see little difficulty in following them, as adapted for the circumstances with the agreement of the regulatory authorities.

Further Information

We appreciate this opportunity to supplement the record on both technical and ethical issues. On February 17, 1993, the Organizing Committee of the HGD Project held an all-day workshop on the ethical issues raised by the Project. We are attaching a copy of the summary and conclusions of that workshop to this testimony to provide the Committee with a resource for further exploration of these issues. In addition, we believe it would be useful to the Committee if we addressed briefly two other issues: the possible racist misuse of the Project's data and findings, and the implications of the HGD Project for public education in genetics.

Racism is an ancient scourge of humanity. We would be deeply disturbed if the Project were to contribute to its survival or spread; in fact, some of those involved in this Project have been in the forefront of fights against earlier efforts to misuse genetics in racist ways. We believe that the Project is likely to undercut conventional notions of race and underscore the common bonds between all humans.

For example, the Project will confirm and publicize that most genes are distributed among human groups in ways very different from skin color or facial or body features. Those characteristics, which are largely though not exclusively driven by climate, tend to be visible, to be homogeneous within ethnic groups, and to be distinctive from other ethnic groups—which is why they form the basis of traditional notions of race. Most genes, which are not directly related to the group's ancestral climate, vary tremendously within ethnic groups and only slightly between ethnic groups. The pygmy populations of Africa are quite different in some outwardly obvious ways from most other human populations. But, over the entire genome, the extent of diversity within any one pygmy village is nearly as great as the diversity within the entire human species. The overall differences between the genes of those pygmies and the residents of nearly any other town or village in the world are likely to be small. Science already teaches us that traditional views of races and of "racial purity" do not make sense; the HGD Project will provide powerful confirmation for that conclusion.

It is nonetheless possible that some findings or data from the HGD Project might be taken out of context in an effort to support racist attitudes. Although misuse of science can never be entirely prevented, the HGD Project takes this possibility very

seriously. We expect to deal with it in part by having "ready response teams" of scientists to respond to controversial claims based on the HGD Project's findings. These scientists would be ready immediately to speak with the media or the public to put such claims into their proper scientific perspective. No one would be censored or denied access to information, but the press and the public would get a balanced assessment of the claim. We know many highly respected scientists from the diverse fields of anthropology, biology, and medicine who we believe would be happy to play such a role. We believe that such an expert and balanced assessment of any controversial findings based on the HGD Project's work will cut against racist interpretations of genetic diversity, not in favor of it.

Whatever the true probability of racist misuse of the HGD's Project's data or findings, concerns about such misuse are both real and legitimate. We intend to reach out to Americans with the greatest reasons for concern—notably Americans from ethnic groups who have traditionally faced racial discrimination—and, through meetings or workshops, to engage in a process of mutual education. We would educate them about the HGD Project and its likely findings; they would educate us about the kinds of misuse that might be expected and the grounds for their concerns.

This leads to the broader issue of public education. Public education is both part of the ethical agenda for the HGD Project and a benefit on its own. Better public understanding of the facts of human genetic diversity will, we are confident, diminish the possibilities of racist misuse of the data. We expect to carry out such an educational campaign.

At the same time, the HGD Project provides a unique opportunity to expand public understanding of genetics. We have all noted the strong interest in this project from non-scientists, from all walks of life. Some are interested in historical questions, some in human evolution, some in their personal family histories. Many of the people who find the HGD Project fascinating have not been as captured by the Human Genome Project. The HGD Project ties into many Americans' interests and imaginations at a more concrete level than mapping and sequencing ever could. The genetics revolution is one of the most (if not the most) important scientific advances of our time. The HGD Project provides a special way to interest more of the public in that revolution and, as a result, educate them about it.

April 13, 1993

To: Participants in the Human Genome Diversity Workshop 3(B) on Ethical and Human-Rights Issues
Non-Participants Who Expressed Interest
Members of the Human Genome Diversity Committee
From: Henry T. Greely, Workshop Chair
Subjects: Summary of the Workshop Proceedings
My Recommendations to the Human Genome Diversity Committee for Structuring the Project's Ethics Component

Attached you will find two documents that have come out of Workshop 3(B) of the Human Genome Diversity Project, on Ethical and Human-Rights Issues, held on February 17, 1993 at the National Institutes of Health at Bethesda, Maryland. As the organizer and moderator of that Workshop, I am sending these documents to all those who participated in the discussions, to those who could not participate but expressed substantial interest in the Workshop, and to all members of the Human Genome Diversity Project Committee, whether or not they attended the Workshop.

The first, and longest, document is the summary of the February 17 workshop. The summary is lengthy and is based largely on tape recordings of the more than five hours of discussion. I have tried to set out, in some detail, the discussions that took place at that meeting, even though the length of the summary probably decreases its audience. The "Conclusions" section near the front of the document can be used as an executive summary.

The second document includes my own recommendations to the Human Genome Diversity Committee on how to structure an ethics component for the Project. Although those recommendations were certainly greatly informed and aided by the Workshop, they should in no way be viewed as reflecting the consensus of the participants.

I want to thank all of you who reviewed a draft of the summary. Particular thanks go to Dr. Irene Eckstrand, whose notes provided an invaluable supplement to the tape, and to Ms. Jean Doble, who along with providing logistical support at Stanford for the Human Genome Diversity Project, also offered her skills as a copy-editor. Apart from corrections of my numerous mistakes in transcription, there were two areas of significant change. First, conclusions number 1 and 11 were toned down in deference to a comment from Dr. Ellis that they seemed to prejudge the issue that he thought needed exploration. I believe I accurately captured the view of the majority of the people who participated, but I was willing to soften the comments in the interests of consensus. Second, Dr. Cavalli-Sforza has expanded his comments from the third session, which fell, in large part, into the gap of our tape recordings.

The funding, speed, and immediate future of the Human Genome Diversity Project remains in some question, but I have no doubt that this work will be done. I believe that the Workshop made a very valuable contribution in exploring at least a few of the ethical and human-rights questions to which this kind of research will

inevitably give birth. And it was fun. I want to extend sincere thanks (on both counts) to all those who took part.

If you have any questions or comments, you can reach me at (415) 723-2517 (office telephone), (415) 725-0253 (office fax), or henry.greely@forsythe.stanford.edu (e mail).

HUMAN GENOME DIVERSITY PROJECT
SUMMARY OF PLANNING WORKSHOP 3(B):
ETHICAL AND HUMAN RIGHTS IMPLICATIONS

TABLE OF CONTENTS

HUMAN GENOME DIVERSITY PROJECT
SUMMARY OF PLANNING WORKSHOP 3(B): ETHICAL AND HUMAN-RIGHTS IMPLICATIONS

The third planning workshop of the Human Genome Diversity Project was held on the campus of the U.S. National Institutes of Health in Bethesda, Maryland, from February 16 through February 18, 1993. The second day of the workshop was devoted to an exploration of the ethical and human-rights implications of the Project. This open meeting centered on three roundtables, involving 12 invited participants, and the resulting discussions among all those present. Attendees and their affiliations are listed in the attached Appendix A. The discussion was guided by a schedule and list of possible issues, distributed to all present and attached as Appendix B. The meeting was organized and chaired by Professor Henry T. Greely, the principal author of this summary.

This is a relatively complete, and thus lengthy, summary of the comments at the meeting. The beginning of the summary sets out as conclusions some issues on which there appeared to be widespread agreement, but those conclusions are not intended to serve as a set of detailed recommendations. The meeting organizer is distributing his recommendations in a separate memorandum; recommendations from others who attended the meeting are welcome and will be distributed by the meeting organizer to the participants and to the Project committee.

Conclusions

These 11 conclusions represent the views of most, if not all, of the individuals invited to participate in the meeting. These are the minimum conclusions to be drawn from the workshop and should not be viewed as precluding any participants from setting forth additional views.

1. The ethical concerns raised by this Project are both real and significant. Although no participant stated that those concerns are so serious that the Project should not go forward, the participants agree that careful attention must be paid in designing and executing the Project in order to minimize the risks of harm.

2. Consideration of ethical issues needs to be integrated into the Project's decisionmaking, both in the planning stages and, on a continuing basis, during the life of the Project.

3. The Project should be designed and executed with help from the populations to be sampled as far as is feasible, although the participants realize that there will often be enormous logistical barriers to such assistance.

4. If any funding is obtained from federal agencies, American laws and policies concerning informed consent must be followed, but applying it in a manner that provides useful information to the populations to be sampled will sometimes prove difficult. No one method of providing informed consent will be

appropriate for every population. This kind of sampling has been done in the past in projects financed by federal agencies, so it seems likely that these difficulties can be surmounted. The Project should collect samples of informed-consent protocols for previously approved research. It may want to create several model protocols or to review itself the informed-consent protocols of researchers who seek to collect samples for it. The structure for central review, in any, of informed-consent protocols requires thought.

5. The Project should consider beginning its sampling with populations that raise the fewest ethical and political problems. It should give special consideration to beginning with populations in countries that sponsor the Project. Experience with those populations, and a record of success in dealing with them, may be very helpful in sampling more vulnerable groups.

6. It is not clear whether any governments of populations to be sampled will seek payments in return for the collection of DNA samples within their borders. The Project should consider what approach it would take to such requests. Because of the many sensitive issues involved in patenting genes, human or otherwise, the Project may want to agree that no genes may be patented using as a basis for the patent any work done on samples collected through the Project. It may also want to agree that populations or countries will receive some form of payment in the unlooked for event that samples collected for the Project lead to products of commercial value. The form of such payments requires further thought.

7. Some people will almost certainly attempt to misuse the Project's data and findings in support of racist or nationalist ends or what Dr. Juengst has termed "demic discrimination." Whether that misuse would have any significant consequences is unclear, but the participants believe the Project has a duty to try to minimize the effects of such misuse.

8. A program of public education would probably be a useful, and perhaps an essential, element in efforts to limit the misuse of Project data and findings.

9. As part of such an education and information program, it is important that the Project define itself, its goals, and its limitations to the public, rather than allow it to be defined by others. Examination of past uses and misuses of scientific attempts at human differentiation may be a useful part of this effort.

10. The Project should ensure that it is informed about the uses to which its samples and data are being put and the conclusions that are being drawn from them. It may want to be able to respond quickly to published work or press inquiries in order to ensure that claims made on the basis of the Project's work are put into their proper scientific context.

11. There is no reason to believe that the ethical concerns raised by this Project are insurmountable. Most of the participants believe that, with appropriate safeguards, the Project should proceed. The Project should, however, evaluate throughout its life the ethical consequences of its work. If

unexpectedly strong and negative effects appear, the Project should be willing, if necessary, to bring itself to an early end. The Project should provide data that may be invaluable in answering important questions in a wide variety of fields, including genetics, anthropology, history, linguistics, and others that cannot yet be guessed. The value of this research and the urgency caused by the continuing disappearance of isolated human populations makes the ethical concerns all the more important. If the Project does not proceed carefully and properly, it could spoil the last good opportunity to obtain some of this data.

Workshop Summary

This workshop session convened at 9:30 a.m., February 17. Prof. Greely welcomed those in attendance, introduced the roundtable participants and the present members of the Human Genome Diversity Project committee, and made a variety of administrative announcements. He laid out his goal for the meeting -- not the resolution of many, or perhaps any, of the ethical and human-rights questions raised by the Project, but instead a narrowing and focusing of the issues involved. Dr. Marcus Feldman, a population biologist at Stanford and a member of the Human Genome Diversity Committee, then briefly explained the Project and described the planning workshops held thus far.

First Roundtable Discussion -- Collection Issues

The first roundtable discussion covered the issues involved in collecting blood, DNA samples, and information from populations of interest to the Project. Prof. Greely noted some of those issues briefly, then turned the meeting over to the four speakers on that issue: Dr. Joan Porter, Dr. Kenneth Weiss, Dr. Rachelle Hollander, and Dr. Gary Ellis.

Dr. Joan Porter

Dr. Porter, senior analyst with the Office of Protection from Research Risks (OPRR) of the Department of Health and Human Services (HHS). The OPRR implements HHS regulations that protect human subjects. She began by reviewing the historical development of ethical strictures on human experimentation, starting with the Nuremberg Code and the Declaration of Helsinki and drawing special attention to the Council for International Organizations for Medical Sciences (CIOMS), which also provides guidance on ethical issues.

In the United States, federal regulatory law is based on the Belmont Report, prepared by the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research in 1979. That Report focuses on three principles:

- Respect for Persons:
  This principle requires researchers to seek informed consent from subjects and to provide special protections for people with diminished ability to consent.

- Beneficence, which includes the principle of non-maleficence:
  Non-maleficence means doing no harm; beneficence means doing good. This principle is realized largely through careful prospective and ongoing risk/benefits analysis by an independent committee.

- Justice:
  This principle means many things, including treating people fairly. Equitable selection of subjects is one way to help ensure justice.

Most of the principles focus on individuals, but justice also concerns populations. The Project lends itself to a "macro-ethical" framework, focusing on communities rather than only on individuals. In considering macro-ethical issues, she has found particularly useful the International Guidelines for the Ethical Review of Epidemiological Studies, prepared by CIOMS in 1991, although there are some differences between those guidelines and American law, particularly on informed consent. She believes the Project raises several particularly interesting macro-ethical questions:

> Will the studies be able to maximize benefits to the community by communicating the results?
>
> Will there be a way to minimize harm by avoiding stigmatization or prejudice or loss of self-esteem or economic loss?
>
> Can harmful publicity be avoided?
>
> Can some kind of confidentiality be retained for groups?

Dr. Porter then set out ten other areas of concern that she believed this group and the Project's organizers must address.

1. What are the real risks and benefits to the individuals and communities? Physical risks are probably minimal, but it is important to anticipate any social risks.
2. Will genetic information be individually identified?
3. Will the sampling process lead to any information of use to the individuals sampled? For example, if a subject's blood had a genetic marker for breast cancer, could (and should) any useful information be conveyed to that individual? What would happen if everyone in the sample had the marker?
4. Will information about disease susceptibility have any possibly subtle effects on employment, immigration, or access to health care for the subjects?
5. Will samples be taken from families and, if so, what concerns are there about pressures to participate, recruitment issues, and problems of confidentiality that have been recognized in large pedigree studies?
6. Will any inferences be made about the relationships between non-human primates and certain groups?
7. Are the groups of most interest to the Project among those most

disenfranchised in their societies and, if so, what are the implications of this status? Who can give permission for such groups?

8. Will DNA from non-living persons be collected? How can community sensitivities about that kind of collection be taken into consideration?
9. What kind of local independent assessment is important? If there are 200 sample areas, will there be 200 institutional review boards (IRBs)? HHS generally prefers local IRBs but perhaps there is an equally effective and more efficient method to do this.
10. How can we ensure appropriate recognition of collaborators, particularly those from other countries?

The Project presents an unprecedented opportunity to advance some areas of human knowledge, making it crucial that we think through the ethical and legal issues carefully.

Dr. Kenneth Weiss

Dr. Weiss is a professor of anthropology at Pennsylvania State University and a member of the Human Genome Diversity Committee. He started by noting that many of the populations to be sampled will be in urban areas. For them, many of the standard answers to ethical issues will apply. Dr. Weiss focused his presentation on the special problems of sampling populations that are living in isolated circumstances and under relatively traditional conditions. He stressed that, because of differing conditions, no one approach would work for every population.

The first challenge will be to explain the Project's goal to these populations. Explaining that the Project wants blood isn't difficult (though it may lead to rejection), but explaining the goals can be harder. For example, he noted that describing one of the goals of the research as discovering information about the origin of a population would be meaningless, or perhaps insulting, to many populations that have their own deeply held beliefs about their origins. Such an explanation could do cultural harm as well as jeopardize the efforts to get permission. In general, explanations must to be made by people who are very familiar with the populations, whether they call themselves anthropologists or not.

All communications must occur in the local language and each must to be tested to make sure it is correctly understood. Although that requirement is not very different from how any kind of epidemiological survey is done, it is more challenging in uncommon languages.

Whether the identity of a studied population should be publicly revealed through the research is an interesting point. Anthropologists have long been divided over the degree to which they should identify specific populations with which they work. This would have to be done differently in different contexts. Identifying isolated communities can sometimes harm those communities and sometimes work to their benefit. This is a real issue, although one without a blanket answer.

The disease questions are interesting, although breast cancer is a bad example for most of these populations. Something like HIV infection is more sensitive because it might have political ramifications. In some places, disclosure of infection could bring help, and in other places, it could bring harmful attention. Someone familiar with the area and the population needs to decide how to deal with samples that show signs of disease before any collections are authorized. No one standard can apply everywhere, but wherever possible, people with current disease should be referred to the national health systems when possible. In many countries, referrals to a national health-care system can make a difference in the health care provided to individual subjects, which may lead to at least a temporary quid pro quo. For a few populations, there will be no health-care system to which they can be referred.

In some populations, payments are required for the research, either to local leaders or to individuals in the studied group. In some of those populations, payments will have to be made to a local leader because paying individuals directly rather than the group leader would disrupt the social system. A person who has experience with that group must decide these issues.

Coercion means different things in different situations. Bravado or lack of bravado, for example, may have a major effect on who volunteers. If coercion happens, it will be from within the population. Sometimes a person who works within the group will have to be persuasive in order to get consent. In many populations, getting a signed consent form will be impossible. People will not sign any form, no matter what it says, because of their intense suspicion that any written form will be used against them.

"Immortalization" can be a very sensitive term and should be avoided when talking about the intended creation of cell lines. (Someone suggested using "transformation," the standard European practice.) Whether to tell people what you intend to do, as a technical matter, is a difficult question. Translating the concepts will be very hard.

As to anonymous samples from dead people, the National Graves Protection and Repatriation Act in the United States provides rules for the return of skeletal remains. If you can identify an existing successor group to which the remains are related, that group's permission is necessary; if you cannot identify an existing group to which the remains are related, he believes they may be used without permission.

The use of fetal and placental tissue will, of course, be governed by local rules and morals. There is no blanket answer to whether you should be able to take samples from people who appear at isolated hospitals outside their home.

Getting permission is complicated and, again, differs in different areas. In working with Native American tribes in the United States, for example, you get permission from the tribal group and go around to the populations with representatives of that group. In some parts of the world, approval must also be sought from a government from outside the population to be studied. In many of

these populations, anthropologists or linguists are already working with the population and can facilitate obtaining permission. Before any collection activity would be funded, its leaders would have to document to the satisfaction of a reviewing group that they understand the permission process in that region and are complying with it. How the review group could check on that information is not clear, but it could not demand an identical process for every population. The Project must always work with people who are familiar with and trusted by the groups to be sampled.

Dr. Gary Ellis

Dr. Ellis has recently become director of the Office of Protection from Research Risk. His office regularly hears from researchers who say that the rules don't fit into their research. His reaction is "let's find a way." Informed consent is a process, not just a form, with the prospective research subject's ability to make a voluntary decision as the key requirement. The procedures should be designed to educate the subjects in language and terms they can understand. The document should be a teaching tool, not a legal document. Lay language, understandable to the potential subjects, must be used even though this is difficult in other cultures. It is important to have a written document so the subject can refer to it in the future.

U.S. federal government regulations lay out eight elements of informed consent.

1. The informed consent must state that the study involves research, must explain the purposes of the research and the expected duration of the subject's participation, and must describe the procedures to be followed with a special description of any procedures to be followed that are experimental.
2. The informed consent must describe any reasonably foreseeable risks or discomforts to the subject.
3. The informed consent must describe any reasonably foreseeable benefits to the subject or others.
4. The informed consent must disclose any appropriate alternative procedures or courses of treatment that may be advantageous to the subject.
5. The informed consent must state the extent to which confidentiality of records identifying the subject will be maintained. If the Project found that someone had a disease and went back and referred them for treatment, obviously the samples would have to be linked to the donors, which may not be the best approach.
6. In research involving no more than minimal risk, the informed consent must explain whether any compensation or medical treatments would be available if injury occurs, what the compensation or medical treatments would be, and where further information about those matters can be obtained.
7. The informed consent must tell a prospective subject whom to contact for answers to questions concerning the research, the research subject's rights, and whom to contact about any research-related injuries. These

questions ordinarily cannot all be answered by the same person for reasons of limited knowledge or conflict of interest.

8. The informed consent must state that participation is voluntary and can be declined without any effect on eligibility to benefits and that participation can be discontinued at any time without any penalty or loss of benefits to which the subject is otherwise entitled. The Project needs to consider what discontinuation means in a context where there will be ongoing cell lines and databases.

Dr. Rachelle Hollander

Dr. Hollander is Director of the Ethics and Values Studies program of the National Science Foundation. She began by noting that the rhetoric of science romanticizes or mythologizes the science. Research subjects, on the other hand, tend to de-mythologize science. Both the public and geneticists need to be educated about the Project. Genetics has a mixed record in its effects on people. The urgency of this Project needs to be moderated by the recognition that it is important to go slowly enough for the public and the scientists to understand the consequences of the science.

The Human Genome Diversity Project was mentioned at the AAAS meeting and was the subject of some misunderstanding. For example, some wondered why we should study diversity when geneticists can create diversity. Others wondered why people were concerned about preserving DNA but not ways of life. There are concerns about evolution within humans because the public and scientists often equate "early" to "primitive" to "less valued."

The need for the involvement and approval of affected communities, from local level to the nation-state, is crucial. What counts as "owning", or "having access," or "benefit" will differ in different cultures. Does this Project involve questions of common human heritage? Will there be questions of ownership or disputes over property rights? How will that be managed? Property rights will make a difference in obtaining informed consent because it affects what you are seeking consent to do. Reciprocity is fundamental and so the Project will have to think about the interactions between the researchers, the sponsoring governments, the populations, and the nation-states that include the populations.

The process of answering these questions must empower individuals from affected populations. They should be involved in developing the Project as well as in allowing access to populations. Informed consent as a mark of respect is very important; respect is an idea shared by many cultures, but sometimes expressed in different ways. Does this consent make sense in terms of the community or person giving the consent? Is there reciprocity? What do *they* think is important? That kind of evaluation is important and it is important that it be ongoing, throughout the Project.

The group might want to devise ethical guidelines for the research, but do so with the participation of the groups involved. This will also lead to the development of ideas about what community payback might be appropriate. The process should also incorporate ongoing evaluation of those ethical guidelines.

The document summarizing the second workshop ends by saying that all materials and data not confidential should be available to all. What that means requires considerable thought.

Discussion

Prof. Greely noted the tension between centralized review for western standards and federal rules, on the one hand, and making allowances for local conditions and cultures, on the other. Devising a process that does both is not a trivial issue.

Dr. Paul said it is useful to note what we know about informed consent in American medicine. We know a lot about how it works, and it doesn't work as "a teaching tool" to help patients learn. For American researchers and doctors, the phrase is "getting consent" and it is viewed as a defense against malpractice suits. Nancy Press's recent research on maternal serum alpha fetal protein screening showed that most of the women signing the consent form either didn't know that they had had the test or thought the test was some form of therapy for the fetus. The true significance of informed consent differs from the ideal even in the United States. (Prof. Greely noted that "consent" has become a different kind of verb in American medicine -- rather than patients consenting to treatment, doctors "consent" the patients.)

Dr. Giddings seconded those remarks. He thinks the Project runs a risk of getting into a political morass. The Project will inevitably have to deal with local political structures, and many people in those structures will not believe whatever the Project says because its representatives will be from the developed world. He hesitates to suggest getting involved with the United Nations bureaucracy, but it might be a good idea. The political problems are so great that the Project must either keep a very low profile or seek political cover, and a low profile is impossible. The Project might want to get involved with the World Health Organization. It might also want to divide the populations of interest into low, medium, and high political risk. It could sample the low risk populations first, as well as populations from the sponsoring nations first. Pursuing the Project poorly might poison the well forever for future scientists. Lots of people will be very suspicious about the Project. Even though those suspicions may be inappropriate, you have to deal with them.

Dr. Siniscalco pointed out that he has a lot of experience on these issues from his work with the European Commission. The study of human genome diversity has been done for a long time -- it constitutes the core of human genetics. Various groups have been doing this kind of research and have been interacting with local governments as well as funding organizations. The main effort should be to advertise the importance of the Project. He suggests something like the DNA Learning Center at Cold Spring Harbor in an effort to educate the public. He exhibited a copy of the Cold Spring Harbor exhibit in a museum in Sardinia and found that the public was extremely interested. He believes the Project also needs to emphasize the crucial importance of environment in its interaction with genes. The Project should not need to keep information about individual samples that identifies their donors, because

individuals will need to be tested only if a medically important gene is found. There are rules and regulations for this kind of research, but the Project should focus on inventing a system to uphold those standards. We have an enormous opportunity; but we need to explain the Project to people in order to gain their trust.

Dr. Reid pointed out that although this kind of work has been going on, this Project will face new obstacles. One reason is the scale of the Project. It will attract attention from people who wouldn't otherwise know that this kind of research is going on. The international political context will also be disturbing. The combination of the Rio discussions and the ethical dimensions raised by dealing with human genes will attract an enormous amount of attention. Education about the motivations for the work won't be enough; it comes down to a question of creating a process that makes this workable. He shares some of Dr. Giddings's skepticism about the United Nations but believes that it deserves a very hard look because of the certifying role a U.N. connection might play. He also thinks it would be very valuable to involve the subjects of the research in the planning, although he recognizes the tremendous logistical problems. Finally, it will require tremendous anthropological expertise in the populations to be sampled; maybe that should be factored into the considerations for choosing populations. (Prof. Greely pointed out that the existence of such expertise had been one of the considerations the second workshop used for suggesting populations to be studied.)

Dr. Juengst agreed that the Project will have a high visibility, which will change the dynamics. That has happened with the Human Genome Project; mapping and sequencing studies had been going on for a long time, but a coordinated Project with a capitalized name becomes a lightning rod. As a result, many of the rules for various studies are being reinvented, at least to provide coordination within the Genome Project. It may be the case that all that is needed here is to codify existing rules for pursuing this kind of work.

Dr. Cavalli-Sforza stated that we may be making the problems seem more difficult than they actually are. He has collected samples from some of the most isolated and difficult locations in the world. You must have the permission of the local authorities, or they may put you in jail, and you must have the permission of the local people, or they won't participate. The permission issues are therefore self-policing. The important issue is the information. Right now, we cannot name any immediate benefits, except possibly with respect to vaccination, which is itself controversial among many people. It is very difficult to explain the science. A traveling Cold Spring Harbor show is a good idea and may work in the cities, but it won't work in the villages where most of the sampling will be done. It may be very difficult to acquire a consent that is as informed as the American regulations seek, but we should go as far as we can.

He has been trying to get the involvement of UNESCO. UNESCO is preferable to WHO, because the Project has no immediate medical purpose. He had a conversation with the Director General of UNESCO, who was very

supportive. He agreed with Dr. Weiss that it was impossible to have general rules; we need to stick to the general rule, "do no harm."

Dr. Schneider was reminded of the fact that the largest samples of blood for blood grouping came from troops on the Salonika front during the First World War. Dr. Hirzfeld, who, with his wife, collected the blood, said in his memoirs that how he collected blood depended on the nationality of the troops. The British troops would do it if they were told it was "for science"; the African troops would do it if told that they could get out of military service; and the French troops would do it enthusiastically if told it would let them know with whom they could make love with impunity. Consent issues go back a long way; scientists are pragmatic and do what they have to do to get the material. He suggested that the proper model might not be medical studies, but traditional anthropological studies because this is not, basically, medical research. Are consent forms used by anthropologists? Another important question is what will be done with the drawn blood. Will this be used as a storehouse for future work, or just be analyzed once? If it is used for a long time, one can't be sure what the uses will be.

Dr. Weiss said that anthropologists don't use forms, but they do get the permission of the groups they study. The Human Genome Diversity committee has known from the beginning that local government approval will be essential and has thought about technology transfer within that context. He thinks we are flattering ourselves to think that the Project will have a high profile. Most of the time, the permission is local, from government officials who want to know that you are there and that you are not going to stir up trouble. Anthropologists get approvals in most parts of the world without much trouble, although there are occasional, usually temporary, exclusions. Some groups like being studied. Groups are very different; confidentiality is meaningless in some contexts. In some groups, doing something in someone's house, away from the group's sight, would be seen as offensive to other members of the group. You more often have to include people you don't want rather than exclude people you do want. To the Committee, from the point of view of funding the Project, it is very important that the Project promises not to reveal individual's names to outsiders.

Dr. Ellis asked who is sponsoring the Project; the response was that we don't know yet. The workshops are planning workshops to put together the framework for writing proposals. Prof. Greely pointed out that if the Project wants U.S. funding, the U.S. regulations will have to be followed for legal reasons, as well as for ethical concerns. He agreed with Dr. Weiss that the application of the principles would necessarily differ from culture to culture. He agreed with Dr. Siniscalco that this kind of work has been going on, so protocols have been approved. Finally, he agreed with Dr. Juengst that, perhaps, "codification" of protocols that have already been used would be sufficient to deal with these issues.

Dr. Ellis added some political advice, saying that the Project should define itself to the public in simple terms first, before someone else does it. It would be a shame, for example, if someone said this Project was intended to find the

genetic basis for violence around the world and the *Washington Post* reported that.

Dr. Giddings said there is a substantial amount of evidence to indicate that this Project is being prosecuted by some guiding intelligence. There will be no way to go forward without a lot of local involvement -- there may be no feasible alternative to funding from something other than a multinational consortium. For political and financial success, this will have to involve multiple governments or international agencies.

Dr. Bodmer pointed out that Europe is involved. The European Community has provided some funding and is interested in providing more. European populations are to be sampled as part of the Project, with at least some European funding. There is interest in making genetic diversity part of the next European funding framework for the biological sciences.

Dr. Schneider asked for clarification about the boundaries of "Europe" for work funded by the European Community. Dr. Bodmer replied that the European side of the Project was able to act outside the boundaries of the European Community, including doing work in Eastern Europe and the former Soviet Union. Dr. Siniscalco noted that the European group is exploring the possibility of Soros Foundation grants for working in Eastern Europe and the former Soviet Union.

Dr. Piazza asked whether it would change things if we collected mouthwash rather than blood. Prof. Greely said it would probably change only the details of the consent, but wouldn't make major changes in the process.

Dr. Feldman noted that throughout the Committee's endeavors, it has had close contact with both China and Japan. The vice president of the Chinese Academy of Sciences was present yesterday and will be tomorrow.

Second Roundtable Discussion -- Payment and Property Issues

The second roundtable discussion explored issues concerning the possible commercial value of the Project's samples, cell lines, or findings. Prof. Greely explained that he had included these issues in the workshop because he believed the recent Biodiversity Treaty would lead some countries with populations of interest to the Project to seek royalties or other payment for sampling the genes of populations within their borders. He then turned the meeting over to the two speakers on that issue: Dr. Val Giddings and Dr. Walter Reid. (The third scheduled speaker, Dr. Jason Clay from Cultural Survival, Inc., was not able to attend because bad weather grounded his flights from Boston. Dr. Clay is receiving copies of this summary and other correspondence concerning the Workshop and has been invited to submit a statement of his views for distribution to the group.)

Dr. Val Giddings

Dr. Giddings, whose training was in human genetics, participated in the U.S. government's negotiating team for the Biodiversity Convention. He read

excerpts from the Convention's provisions concerning technology transfer and payment.

He said that the Convention and the negotiations from which it emerged will have important implications for this Project. It is a new era for genetic resources. The Food and Agriculture Organization's principle of "free access" to genetic resources is not going to survive. Movements of genetic resources across international boundaries in the future will be subject to a variety of constraints, particularly if there is any reasonable inference that the resources could result in commercial products. He urged that the Project must anticipate those constraints and build in some kind of agreement for royalty partitioning or some other provision for distribution of fair and reasonable benefits. Otherwise, it won't be able to get permission to take samples.

Dr. Walter Reid

Dr. Reid is associated with the World Resources Institute and has a special interest in issues of global biodiversity. At Rio, Dr. Reid participated with other organizations in developing a strategy for conserving biodiversity around the world. That process was not as polarized as the Convention negotiations, but it also reached the conclusion that the era of treating biodiversity as a common heritage or open access resource is over. Arguably, the distribution of benefits from the old regime was not equitable, and developing countries see opportunities to get more benefits in the future, through both financial arrangements and technology transfer.

In the Convention negotiations, intellectual-property rights and technology transfer were most contentious. This Project will not have a profile as high as the Convention's, but he suspects that the Project could not go into countries to get permission for sampling without encountering officials who were familiar with the Convention. Those officials might draw on the Convention's terms to deal with the Project's request, particularly in countries that signed the Convention.

He drew a parallel to the existing seed banks, which were collected under the common-heritage regime. For example, there is a bank of 60,00 or 70,000 varieties of rice in the Philippines, with copies in the United States. The Convention negotiators decided they couldn't deal with the existing collections and so excluded that material collected in trust for humanity. How to deal with those materials remains a subject of discussion. Some seed banks are talking about patenting their materials to make them available; others prefer a situation whereby the seed banks only transfer material subject to a materials transfer agreement that includes a provision that no one seek a patent on the material.

Because this Project seeks discoveries for all humanity and commercial profit is not its main concern, he suggested that the Project sign material-transfer agreements with the host country and with anyone who took samples from the Project. Those agreements should provide that no genes in the materials would be patented. Although such a ban on patents might not have much practical importance, the issue of patenting genes is politically very sensitive in many countries. A stipulation that no one would patent genes taken from cell lines from

the Project, combined with an undertaking to share profits if the Project's cell lines led to products with commercial value, would be very useful. He argued that unless the Project made a preemptive strike on the patent issue, it would open a Pandora's box.

Discussion

After an inquiry from Prof. Greely, Drs. Giddings and Reid said that the Biodiversity Convention neither expressly includes nor expressly excludes coverage of human genetic resources. Dr. Giddings agreed with Dr. Reid on the importance of a preemptive position on this issue, though he would allow more scope for intellectual property protection for materials that had been improved in accordance with international common law or U.S. patent law.

Prof. Greely, seeking clarification, posed a hypothetical question. If the Project discovered that an isolated population had a gene that protected it against an infectious disease and a researcher using the Project's cell lines isolated the gene, found the protein for which it coded, and patented not the gene but the protein as a basis for a drug, would Dr. Reid expect problems? Dr. Reid said that the major political concern in most countries was the actual patenting of the gene. Protection for the protein would be appropriate, although he thought it should come with some sort of return to the host country. Dr. Reid thought the discovery of a human gene that a commercial firm wanted to add to a crop's genome would be a much trickier issue, although one unlikely to arise. In such a case, the firm would want to patent the gene itself. It would be useful to the Project to say, from the beginning, that no patents will be sought for genes taken from the Project's samples. A firm interested in patenting the gene could always go back independently to the country where the gene was found and make whatever financial arrangements were appropriate, but that would not implicate the Project.

Dr. Giddings said limiting a ban on patenting to the gene itself would eliminate many, if not all, of the problems he had seen with Dr. Reid's position. Even so, he believes Dr. Reid's proposition would meet with opposition from the governments of some developing countries. The developing countries often have a completely erroneous view of the process of product development and one that will often hurt innovation. Questions of intellectual-property rights should not be involved significantly in this Project. This Project is a research project, not intended to have commercial value, and should deal with these issues only to the extent necessary to avoid real political problems.

Dr. Kenneth Weiss said the Committee wanted to make sure that anyone, anywhere in the world, can get access to the samples for research. He also pointed out that none of the property rights, or benefits, involved here are likely to get back to the villages. He argued that we aren't really talking about benefits to the people who would be sampled. Prof. Greely said he hoped they would discuss how to try to get the funds back down to the sampled population. Dr. Giddings said the Food and Agriculture Organization had been dealing with this at some length and found that there was no good solution. He believes that the Project would be better off avoiding this very tricky issue.

Dr. Paul noted her disagreement with Dr. Giddings's view of intellectual property and the developing world and then asked for clarification of the positions of Drs. Giddings and Reid. Dr. Reid said that the issue of patenting genes may be relatively trivial as a practical matter, but that possibility would raise a lot of political concerns. Unlike Dr. Giddings, Dr. Reid thinks developing countries, such as India, might well accept his formulation. For one thing, it would allow them to continue to sidestep the tricky issue of the patentability of human genetic material.

Dr. Schneider asked whether the Project could proceed by excluding any possible commercial use. Prof. Greely said it would be hard to exclude commercial use while keeping open access to the cell lines. And he added that the distinction between research use and the subsequent commercial use would be difficult. Dr. Schneider said perhaps you could prohibit the commercial use of the sample; it was pointed out that copying the materials would be easy.

Dr. Siniscalco raised the possibility of having any royalties or profits derived from the Project going to an international body to be used for general human benefit. Prof. Greely replied that the question was really one of what host governments would accept. Dr. Feldman pointed out that if part of the funding came from U.S. government sources, researchers would have to sign a standard patent form. He was not sure what effects, if any, that form might have on obligations to the host countries. No one present knew the answer to that question, although Dr. Eckstrand pointed out that almost every bilateral agreement involving the NIH contains annexes dealing with the division of property rights in different countries. These generally reserve intellectual-property rights in different parts of the world to participants from different countries.

Dr. Reid noted that one could have any royalties go to a UNESCO fund to be used to promote technology transfer to developing countries. UNESCO management would alleviate many political concerns; technology transfer was a major concern in the Biodiversity Convention negotiations. Prof. Greely noted that this still would not directly help the sampled populations, although Dr. Weiss pointed out that about three-quarters of the populations to be sampled would be groups that were integrated into their country's economy and culture.

Dr. Reid said there had been a great increase in awareness among ethnopharmacologists of the need to help the communities involved in their research. There are not formal contracts, yet, but there has been a great increase in the benefits going back to the affected communities. The New York Botanical Gardens has been funding health centers and other benefits the populations want. This has been going on without the attention of national governments.

Dr. Evans pointed out that the significance of a gene can only be seen from the entire samples; one cannot know that one population's allele is important without comparing it to the entire sample. The complexity of

compensating the populations from all over the world is daunting. She also pointed out that this Project should not be viewed in isolation, but as an offshoot, to some extent, of the Human Genome Project. In that Project, the thrust of the discussion is against patenting genes but in favor of patenting their useful applications. Commercial interests will come in at that stage. If the information is available to industrialists, you cannot stop them from trying to use it or to protect any valuable applications they make. She pointed out that patenting genes is different from patenting useful knowledge derived from them.

Prof. Greely said he thought Dr. Evans's point on the importance of all populations was very important. He recalled a suggestion Dr. Eric Lander had made to him in a telephone conversation that any royalties should be viewed as a "mutual fund" for the benefit of all the sampled populations. This, he felt, plays back into the idea of UNESCO or another international organization as fundholder to promote the interests of indigenous peoples or, perhaps, those who contributed samples to the Project.

Dr. Paul noted that the discussion was confusing two separate questions -- who is the appropriate fundholder and what is generating the funds. There will be disagreements between proponents of intellectual-property rights and proponents of the Third World's interests. The Project will have to confront those broader issues and not just questions of the patenting regime.

Prof. Greely urged that the Project needed a consistent approach to avoid having different arrangements with different countries.

Dr. Siniscalco pointed out that patenting a gene, gained through one method, could be circumvented easily. Those who gain will be those who build the product most efficiently; another slightly different version of the gene could always be found for use. Prof. Greely argued that, even if the practical implications of intellectual-property rights in these genes were meaningless, it may be necessary for political reasons for the Project to have a position on this point. Dr. Giddings suggested that it might be easy for the Project to offer high royalties, because the Project would not, in fact, be conceding much of value.

Several people disagreed over what legal or ethical entitlements might be owed to those people or populations who provided samples that led to commercial products. The *Moore* case in California was evidence that at least one donor thought he had property rights to his cells and that issue has been settled, as far as those present knew, only in California. Dr. Giddings urged that the Project need not answer all these issues before going forward; he supported Dr. Reid's general approach as a good way to deal with the issues.

Prof. Greely noted that some ethicists felt that human genes, and humans, could not ethically be turned into market commodities. He did not know whether anyone at the workshop took that position but sought comment on it. Dr. Evans said it was very important and noted the different European laws on payment for blood donation. Prof. Greely noted American laws banning payment for most organs, and Dr. Evans said the same was true for many countries in Europe.

Dr. Evans noted that the European Project was considering seriously advertising for donors so that everyone who was sampled would have made an affirmative decision to come forward and participate in this Project. She recognized that this method could not be extended to every population, but thought it should be kept in mind. The information seeking participation could include language dealing with property issues.

Dr. Weiss pointed out that this discussion fed back into informed consent. It would be very difficult to explain to people that their blood could be "sold" or, having explained it, to get their permission to take samples. Distributing the subjects' genetic material as a business would raise great concerns. Dr. Feldman pointed out that one goal was to have several cell banks, located in various regions. The laws of distribution of the cells are likely to differ in major ways among these countries; whatever the Project wants to do may be overriden by those local laws.

Dr. Schneider noted that he personally would be much less likely to donate blood if he were told it would be used to make a profit. He then raised the question of whether any governments were very eager to participate in the Project and even wanted to do a more thorough study. Dr. Evans said that France had already done a province-by-province study of its populations and was, as a result, more eager to participate in the overall Project.

Dr. Weiss noted that some countries might demand, as a short-term quid pro quo, laboratory facilities in return for giving permission to sample.

The U.S. government is interested in looking at variation within its borders for forensic purposes. That raises concern among some groups. Dr. Cavalli-Sforza pointed out that the Project was not likely to have major value for forensic purposes because most of the populations it will sample are not heavily represented in the U.S. population.

Dr. Bodmer noted that there had been a Japanese study of 128 different populations. The cell lines from that study were considered public domain, with no identification of individuals. There was no government involvement in this because the people who collected the data were from the countries where the sampled populations were located.

Prof. Greely noted again his belief that the higher profile of the Project and the negotiation of the Biodiversity Convention would make past experience a poor guide to predicting host governments' demands in the future.

Third Roundtable Discussion -- Racism and Other Possible Misuses of the Project's Data

The third roundtable discussion concerned the possible misuse of the Project's samples or results. Prof. Greely started by saying that he thought the issues in this roundtable were the most important to be discussed today.

Although a variety of misuses is possible, Prof. Greely noted that, as an American familiar with America's historical problems, he considered white-black racism the greatest concern. He then turned the meeting over to the speakers on that issue: Dr. Diane Paul, Dr. William Schneider, Dr. Eric Juengst, Dr. Luca Cavalli-Sforza, and Dr. Robert Murray. (Dr. Murray was detained by an emergency and arrived late in the discussion of this issue.)

Dr. Diane Paul

Dr. Diane Paul directs the Program in Science, Technology, and Values at the University of Massachusetts at Boston and has studied the social history of genetics. She started by saying she thought the issue of racism would be the least important or difficult ethical problem with the Project. Everyone now says they are against racism, including, these days, every racist. We know where we come out on general principles. Difficult issues remain, but they are of a different nature from the discussions we just had over intellectual property. She urged that the Project consider intellectual property issues from a variety of perspectives, including the perspectives of the developing world and its advocates.

She said the Project is likely to reinforce conventional understandings of race and ethnicity. For example, in current genetic screening programs, although geneticists and anthropologists emphasize that traditional race is arbitrary, screening programs use traditional ethnic and racial categories for the screening. This tends to make those categories *seem* more real. She could propose no solution, but it is one way in which the Project will have social implications.

By contrast, she thinks that the specific results of the Project will neither reinforce nor undermine racism, but instead are likely to have very little effect on such attitudes. Dr. Mary-Claire King has said that the Project will undermine racism by showing how much alike we are; Dr. Paul said this is naive because the findings will not be clear-cut. The findings will be a set of statistics that will have to be interpreted to the general population. The discussion will be in terms of "more or less" -- is this is a substantial or insubstantial difference? It is like the nature/nurture controversy. Even if scientists put a number on the heritability of a trait, the number will not settle anything because some people will interpret the same result as large or small depending on their social interests. If the average human being is heterozygous at 10 percent of loci or if a population is polymorphic at one-third of loci, is that a lot or a little? It depends on the background assumptions brought to the question. She would not put too much emphasis on the consequences of the Project's results because, whatever they are, they will be controversial and will be invoked on behalf of every possible claim.

She pointed out that it also is not easy to say what findings support the "progressive" side. In the classical balance controversy, which was a controversy about the extent of human diversity, H.J. Muller argued in the 1950s and 1960s that we are fundamentally alike. At the same time, Theodosius Dobzhansky argued that we were all very different, that genetic diversity was good for individuals and was good for populations, and therefore that we should expect to

find diversity maintained. For Muller it made sense to talk about a normal, wild type of a gene, and for Dobzhansky it was nonsense.

Today we think that if the Project shows that people are all basically alike, that will be good and will undermine racism. But much of the bitter dispute between Muller and Dobzhansky flowed from the fact that they thought different social policies flowed from their views. In that dispute, Muller was the eugenicist and Dobzhansky was the anti-eugenicist. Dobzhansky wanted to show we were different in order to undermine Muller's eugenics policies. Muller's policies made sense if one thought there was genetic uniformity, because then one could conclude nature was striving for a best type. Dobzhansky wanted to say that nature loves diversity and people should not reduce diversity through eugenics. These policy implications are tremendously plastic, and the particular connections people make are highly contingent. As a result, Dr. Paul would not put much hope in the Project's undermining racism. Look at the former Yugoslavia; the Human Genome Diversity Project will probably not have much effect on that situation.

Dr. William Schneider

Dr. Schneider is chair of the Department of History at Indiana University-Purdue University at Indianapolis and has studied the history of research examining blood groups and populations. He began by saying that the Project's results will be misused, but it was not clear by whom. He noted that his analysis is based on his study of the history of similar efforts and not on a deep knowledge of current science. He did note, however, that the Project probably will not produce the results that are anticipated, because, historically, scientific projects rarely do produce the expected results.

Dr. Schneider has studied the use of blood groups from 1900 to 1950 and, in particular, between the two world wars. He has looked at the work done by scientists from the discovery of blood groups and their inheritance according to Mendelian laws and primarily the discovery, during World War I, that blood groups were differently distributed in different ethnic groups. Hundreds of researchers published over 1200 articles in dozens of journals, analyzing tests of several hundred thousands of subjects in scores of countries and colonies around the world. As a result, this effort may have been larger than the proposed Human Genome Diversity Project. The most frequently asked question was whether blood groups could provide a different definition of race. Researchers were also interested in questions of links to disease, insanity, criminality, and so on. The working title of Dr. Schneider's project is "The First Genetic Marker"; the research he examines has both similarities to and differences from the Human Genome Diversity Project.

The first similarity is that the blood group effort was based on a new scientific discovery, in its time probably as revolutionary as the genetics revolution. Second, blood grouping had immediate practical applications in blood transfusions. Third, the blood-group researchers wanted to explore differences between peoples. There had been earlier efforts, but, like the Project, the blood-

group researchers used new technology. The blood-group researchersers were also interested in the origins, movements, and mixings of populations.

The blood-group research began with individual researchers finding samples from whatever groups they could, taking their samples, and publishing their results. Hospitals, soldiers, prisons, and mental institutions provided convenient sources for samples, as did the patients of doctors or missionaries stationed in remote parts of the world. The general model was that individual researchers did their work and published it in journals. Periodically, someone would publish compilations and analysis of that data. Some countries put forth systematic proposals similar to the Project. One of the most ambitious proposals was made in Germany in 1926, by the German Society of Blood Group Research. This Society, founded by a particularly "volkish" anthropologist named Otto Reche and a Navy doctor named Paul Stefan, divided the German world (including Austria) into 900 districts and called for the testing of 500 subjects in each district. It appealed for the cooperation of local doctors through publications in medical journals and in general-interest magazines. After about 12 studies, some based on school children, the Society asked the Prussian state welfare ministry to perform blood testing in all school districts. After the lengthy hearings, the council recommended only limited testing. The Society did not get what it wanted but it did get some legitimacy. It went on to create a journal to publish the results of the work of others.

This interest was not limited to Germany. The Soviet Union created two research centers, one in Leningrad, which urged doctors to send samples to a central location for analysis under controlled conditions, and one in the Ukraine that followed the German model of asking doctors to do their testing on site. Two smaller countries, Holland and Denmark, persuaded their governments to support systematic studies of blood groups within their countries. Both of them used centralized analysis. Between 1919 and 1939, 75 countries and colonies were sampled with over 200,000 subjects.

Was this used for racist purposes? Yes, in two ways. The first came from Hirzfeld's findings that a higher percentage of Type A blood was in northern and western Europe, with increasing proportions of Type B blood in central and southern Europe and moving into Asia. He devised something he called the biochemical index of race, which was simply a population's percentage of Type A blood divided by its percentage of Type B blood. Consciously or subconsciously, this meant that the highest numbers applied to northern and western Europe. It was fairly quickly recognized that the index did not work as a guide to conventional views of race, but studies through the late 1930s continued to use the Hirzfeld index to put the populations of the world into a hierarchy.

In Germany, especially, the Aryan ideologists liked to show charts of the world based on the Hirzfeld index, interpreting the results as an invasion of Europe by Type B blood with Germany as a bulwark against this invasion.

Similar kinds of studies were done in other countries as well. Gypsies, Jews, the Lapps, and Native Americans elicited particular interest.

Dr. Schneider asked whether this story has a moral. The research into blood groups and ethnicity did not represent a totally unmitigated disaster. The studies showed that the simplistic idea that a single number could define race was untenable. The blood groups ultimately provided a body of information that became a fundamental part for the whole field of human population genetics after World War II. The concept of race was different in 1949 than in 1919, at least in part because of this research. Thus, the research did make a difference, at least among scientists. This underscores the importance of access to information. It is a fundamental tenet of science that open access helps determine whether the truth will come out. You need to think long and hard before restricting any access to data.

Dr. Eric Juengst

Dr. Eric Juengst has a Ph.D. in philosophy and directs the NIH program on the Ethical, Legal, and Social Implications of the Human Genome Project. He began by saying that he is an optimist about the ability of the scientific and genetic community to anticipate and to manage the impact of its research. He sees as success stories for such active management both the recombinant-DNA debate in the 1970s and the procedures that came out of it and the discussion of gene therapy and the procedures that were developed for its assessment. He hopes that in the 1990s the Human Genome Project and its efforts to build in assessment will be a success story. He got involved in the Project because he thought we could anticipate the issues and prevent overly deterministic and reductionistic interpretations of personal genetic information and thus prevent "genetic discrimination."

He believes the Diversity Project will take us to the next level of difficulty. The Diversity Project stands to inherit from the Human Genome Project the role of the lightning rod of genetics. The main social risk of the Human Genome Diversity Project is what he would call "demic" discrimination. This is not necessarily racism or racial discrimination, because "racial" categories look quite outdated, but discrimination against particular demes or subpopulations of the human community as a result of the misinterpretation or the misuse of the conclusions of studies done with the data collected by the Diversity Project. He sees three reasons to think the Project should anticipate such demic discrimination.

First, unless it is very careful with its educational efforts, the Project is likely to be perceived by the public as an effort designed to establish a taxonomy of human types and categories. That will inevitably bolster contemporary notions of race and ethnicity. The statement from the first workshop report that the plan is to identify "the most representative descendants of ancestral human populations" worldwide and then preserve the genetic sequences suggests a typology. The counterargument is that the Project will find more similarity that difference but this is called the Human Genome *Diversity* Project, not the Human Kinship Project or the Human Family Project, because the scientifically interesting parts are the differences.

At the very least, a massive educational Project like the one planned for Europe is probably a prerequisite for the Project as a whole. That means the Project will necessarily have a high profile; there will be no chance to do this quietly. Vital issues are how the Project will be perceived and interpreted and whether it be perceived as a typology of human types that begins to look like a search for the "pure strains" of various types of humanity.

Second, the findings the Project is expecting could be used to fuel existing human antagonisms. At the top of the list of the scientific questions are questions about relations to neighboring groups and questions of migrations -- when groups arrived in particular areas. But, he noted, lineage and land tenure are probably humanity's favorite excuses for making claims about social privilege. New scientific evidence bolstering the claims of one party or another could easily fuel existing fights. One can envision two scenarios. In one, the dominant group in a society is included in the Project, has its DNA hallmarked, and uses those hallmarks as inclusion criteria -- if you want to be part of the ruling group, you need to have the right DNA marking. The other side could be the flip side -- a dominant group uses DNA marking to identify groups for oppression. These groups would have to have access to appropriate DNA types or hallmarks to plug into a system of forensic DNA typing, but that will probably coming anyway, given police interest in this methodology.

Third, one reason this Project is urgent is that many of the populations are vanishing. Why are they vanishing? Some are being assimilated into larger populations for no particularly nefarious reasons, but others are vanishing because they are at political or social or economic disadvantage and can no longer maintain the cohesion of the communities that once sustained them. That suggests that some of these demes already face political risks. Dr. Juengst likes the approach of starting with the least politically risky groups, in order to learn from experience what problems to anticipate and how to avoid them.

What can we do to minimize these dangers? Dr. Juengst had two suggestions. First, the Project should be preceded by an educational campaign to define the Project to the public. It would be very dangerous to allow others to say what the Project is about; it would be much better for the Project to take an active role in presenting itself to the public.

Second, he suggested that the Project should include a standing advisory group, akin to an IRB, to review requests for access to information. This group would ask researchers why they wanted the data or samples and what they intended to do with them. Such a group could serve to limit the misuse of the Project's data.

Dr. Luca Cavalli-Sforza

Dr. Luca Cavalli-Sforza is a geneticist at Stanford and a founding member of the Human Genome Diversity Committee. He began by defining racism as a belief in the biological superiority of some group. He noted that there is no biological basis for this belief, either in terms of genotype or in terms of phenotype. To the extent arguments are made for one group's superiority, they

are always based on phenotype and mainly on behavioral traits about which we have very little genetic data.

Questions of the genetic basis of IQ became controversial about at the same time Dr. Cavalli-Sforza arrived in the United States in 1971. Arthur Jensen, encouraged by physicist William Shockley, argued that IQ had a substantial genetic basis and that the American black population had lower IQ's than the white population for genetic reasons. In fact, this argument was based on bad science in a variety of ways, and scientists, including himself, exposed its shortcomings.

The traditional concept of race is not biologically meaningful. There is usually as much or more genetic diversity among the inhabitants of one isolated village as there is between population groups. If we want to understand something about human anthropology or history, we need to look at hundreds of traits because differences between groups are very small and otherwise comparisons are statistically inadequate. This DNA work will only confirm what we already know to be true about the meaninglessness of race. It is extremely likely in fact that this project will reinforce the current notion, obtained from thousands of studies of immunological and protein-electrophoretic variations, that there are no discontinuities among human groups that justify the concept of race. It should also be emphasized that when we speak of diversity we refer to *diversity among individuals, but* it would be silly to limit the study to individuals of a single, narrow group. We already know that when we do so we obtain results that are biased. Unfortunately, most existing data aware not ideal for evolutionary studies because they have been made with genetic variants detected in people of European origin. The study of diversity among individuals should be done on an appropriate sample taken from the whole world. By choosing it judiciously we can obtain information of enormous historical importance, which will be soon destroyed if we do not proceed before economic development cancels the residual differences among ethnic groups.

He believes most people are convinced there is a true discontinuity among ethnic groups, thus justifying the idea of races, and that they derive this belief from the fact that we are all the time confronted by conspicuous differences in skin color and body and face shapes. These traits tend to be reasonably homogenous among individuals of one ethnic group, and different from those of another group. Such traits, however, are far from representing a random sample of genes. They most probably are due to an adaptation to different climatic conditions, made necessary by the evolutionary history of humans. The average genetic trait does not show the same homogeneity within groups and differences among groups as skin color and other external traits known to be influenced by climate.

He agreed that an educational effort is a good idea. He noted that the Project will not add much to the forensic uses of DNA technology because the groups to be sampled are not generally of forensic importance in the countries that are pursuing forensic technology.

Discussion

Prof. Greely opened the discussion by asking how can we minimize the misuse of the Project and whether the concerns are so great that the Project should not proceed.

Dr. Paul said that she had little confidence in the ability of educational campaigns to improve public understanding. Dr. Siniscalco stressed the importance of the Project for "euphenics," improving the phenotype of humans through a better understanding of the interactions of genes and environments.

Dr. Kenneth Weiss pointed out that racism is much older than science and is based in differentiating one group from another. Racists will always find data to misuse, whether or not it comes from the Human Genome Diversity Project.

Dr. Bodmer asked what group would be appropriate to oversee the use of the Project's data.

Dr. Mark Weiss noted that there is plenty of data in the published literature that could be used for racist views. If someone wanted to use the Project's cell lines to generate new data, he would need a fairly complicated lab. Dr. Weiss also noted that there are plenty of examples of racism in modern America, sometimes from surprising directions, and pointed to some aspects of the Afrocentric movement as an example.

Dr. Feldman noted the distinction between misusing bad data, as Jensen did, and misusing techniques in manipulating good data. The markers that geneticists use today are racially biased because they are drawn from white populations, thereby making them, for many purposes, incomplete. He also noted that data can be abused by omitting some data, as in one study that purported to show a high degree of heritability on the basis of studies of 44 pairs of twins, while omitting data on another 111 twin pairs.

Dr. Hollander stated that the Project does have an ethical obligation to seek to limit misuse. She noted that reactions to the Project vary in part with individuals' different levels of hope and fear. She would make a very different prediction from Dr. Cavalli-Sforza, but we cannot know in advance what the effects will be. How can we address problems when we can't know their full dimensions?

Dr. Paul noted that Dr. Juengst had suggested the Genome Project as a model, but she argued that it is a flawed model. With the Genome Project, the scientific work goes forward and the Ethical, Legal, and Social Implications program goes forward, but the conclusions of the ethical, legal, and social implications program affect the science, if at all, in ways that are highly mediated and perhaps not very effective. If the Project is going to take these issues seriously, it needs to set up some kind of forum that is involved in making decisions on how to proceed on the technical side. If the ethical concerns are not integrated with the technical decision, they will not be effectively addressed.

Dr. Hollander noted that Dr. Juengst had earlier argued for involving the affected communities. This involvement needs to be addressed not only in the design of the project, but in some kind of continuing oversight, monitoring, and reconnaissance effort. If for no other reason, having community representatives involved will prepare the Project to have these conversations again and again, as it explains itself to the world.

Dr. Paul said the Project needs to leave open the question of calling a halt in case the problems turn out to be worse than anticipated. The second question in section three of the agenda for this meeting was whether the concerns about such misuse are so great that the Project should not be undertaken. She said that the discussion had taken for granted that the answer to that is "no," but that she believes it should be an open question. Prof. Greely noted that he had put the question on the table and that there had not been any effort to "hide" that question; calling an open workshop at NIH is not consistent with trying to avoid the question.

Dr. Eckstrand pointed out that NIH has a policy of openly sharing resources regardless of what those resources are. The example she knows best is GenBank, the database for all the DNA sequences that have ever been published. It is available to anyone who wants to tap into it. GenBank is maintained by the scientific community, and people have the right to use or misuse the data. By and large, the misuses have been honest scientific ones. She cannot imagine NIH saying it should have a resource of this nature and not make it available to anyone who wants to use it.

Dr. Kenneth Weiss said first that he knows people from American minority groups who are offended that the Human Genome Project is using their tax dollars but is not including them. African-Americans know that a lot of medical genetics cannot be done for them because we don't know enough about their genotypes. If it were decided that it was not important to sample the rest of the world, it would be saying, in effect, that nothing out there is different or important. We know, he said, that this is not true.

Second, before the second planning workshop, which he organized, he had expected the Project to produce a great deal of controversy among anthropologists, who as a group are very sensitive to these problems. Instead, there was almost universal enthusiasm for this. And for biomedical reasons, given the acceptance of the infinite-alleles model that most mutations are unique, there is no excuse for not sampling everyone. He said we know that mutations are going to be different in different peoples. Some of the populations the Project should sample represent tens of millions of people in the United States alone.

Dr. Juengst said he thought for the Project to have this conversation in public, with the press in the room, was a great beginning. More generally, he noted that no one present had criticized or dismissed the reasons to do this Project. Instead, the discussion has centered on how to execute the Project without causing more trouble than necessary.

Dr. Cavalli-Sforza said he was glad that the NIH policy is to have everything open. That is how science should proceed. He then noted that data on some 250 genes have already been collected by classical (non-molecular) techniques. That information is not as good as the data at the DNA level sought by the Project, but it has been collected from a great number of populations. These data have shown that genetic differences between races are small, and perceptible only by statistical investigations. There is no reason to think that analysis at the DNA level will change the picture; in fact, we already have enough information to be confident that it will not. What harm can come out of data of the new type? A trait like IQ is a product of perhaps 200 genes, which are likely to interact in complicated ways and, in addition to react to a poorly known set of external, non-genetic factors and events. It is difficult to believe that a satisfactory analysis of such complicated situations regarding behavioral traits will be possible in the near future, perhaps not even in time for the younger people in this audience to see.

Dr. Murray then arrived and, after a break, spoke.

Dr. Robert Murray

Dr. Robert Murray is a physician and a faculty member in clinical genetics at Howard University. He has long worked with issues of genetic diversity and genetic screening in African-American populations. He began by noting that he had been involved in these kinds of issues for some time. Howard University has been collecting samples and information on genetic diversity among African-Americans. His experience has made him concerned about the possible misuse of genetic information concerning populations with respect to behavioral or social issues. Thus, for example, people might use information about a genetic propensity to low birth weight in one African population to argue that the problem with low-birth-weight African-American infants was "genetic" and should not be used to justify improving living conditions or medical care.

Dr. Murray said he had just come from the AAAS meeting earlier in the week. At that meeting, the media were interested only in the question of crime and heredity, even though there is no data to support a true genetic relationship. He had been interviewed three times, not to discuss genetic screening, which he knows something about, or genetic engineering, for which there is some information, but about crime and heredity, where there is nothing to discuss. A person who misuses the Project's data will have a ready platform unless there is a quick response from someone the press can turn to for information.

Concerns about the social uses of information on genetic diversity arose because of the pressure in the black community to do a better job of collecting bone marrow and kidneys for transplantation. Many people expressed the worry that the existence of different markers would be used to suggest that "these people" should be kept separate.

Dr. Murray said the risk of harm comes from when studies are released and how they are reported. It is important that data be reported only when researchers have a substantial amount of information. For example, in the past few years, highly publicized studies have reported genetic linkages for manic-

depressive syndrome, alcoholism, and other traits. Those linkages were then disputed. Word of the initial studies is widely distributed, but the follow-up that the linkages may not exist gets much less publicity. As a result, people continue to believe, erroneously, that these linkages have been proven to exist.

Our society talks about populations with different backgrounds and characterize certain life styles as "primitive" or "underdeveloped" when those populations live lives very different from ours. If significant genetic differences turn up in markers for these populations, some will attempt to attribute their cultural status to a genetic basis. Wilson's views of sociobiology are believed by many people who think much of our behavior is genetically controlled.

Dr. Murray said he hoped that, in the design of the Project and whatever standards are set for the reporting of the data, significant thought will be given to providing responses or rebuttals to people who draw unjustified conclusions from the Project's data.

Discussion

Prof. Greely noted that Dr. Juengst had earlier suggested something like an IRB to see who should have access to data. He noted that Dr. Murray's concerns might be allayed by an entity that would have advance warning of what was being explored and perhaps even what results were to be published. This entity could constitute a "ready-response team" to try to put findings into perspective.

Dr. Murray said that this had been tried with gene therapy. The human gene therapy subcommittee spent a lot of time drawing up guidelines, including some concerning dealing with the media. They were primarily concerned about early reports' giving false hope if, for example, someone gets evidence of gene transfer in one particular case.

Dr. Siniscalco reminded the group that this Project is the best opportunity to collect this kind of data. Knowledge may be misused but we know that ignorance is at the basis of every misconception. The Project, he said, will give us a chance to understand not just history but the relationship between genetic makeup and environmental influences. In Sardinia, within the space of a few hundred miles, there were major ecological differences and human populations with substantial inbreeding. This has led to huge genetic differences based on the degree of exposure to malaria. The misconceptions of people like Dr. Jensen will draw forth responses from people like Dr. Cavalli-Sforza to correct them. We need to record these differences before human intermingling makes it too late. We must get the data; issues of the possible misuse of the information will be dealt with as it is interpreted. Otherwise, we will not be able to accumulate the knowledge that could benefit everyone.

Dr. Porter noted that the discussion has brought out issues of how much should be centralized and decentralized. Dr. Juengst's proposed oversight body would be a centralized effort to ensure the highest standards of scientific review of the work that will be done on a decentralized basis. Centralization also will help in informing the press and in dealing with the political process. There will

necessarily be some disadvantages to becoming too centralized, including possible stifling of academic freedom or new uses for the data, but the question of centralization or decentralization is one useful way to think about these issues. What kind of oversight body would the Project want to construct to get the work going and to keep it going?

Dr. Kenneth Weiss said that all that the Committee is proposing so far is to organize the collection of data, not to do any kind of analysis of that data, except to run a standard set of markers on all samples collected worldwide. The Committee has not proposed that the Project should be responsible for administering grant applications for analysis. What happens after sampling is a question the Committee has not considered, although it is a valid one.

Dr. Murray said the problem is that collecting the samples will provide a source for abusive kinds of analysis. He had not thought there was any question about the Project's going forward. He thinks it should go forward -- on the whole, it will be beneficial rather than harmful. However, just one or two instances of misuse could destroy all the good the Project will do. The University of Maryland debacle undercuts all the good work geneticists have tried to do. He does not want that to happen with this Project. The discussions today are about trying to design the Project in a way to minimize the possibility of misuse of the data. The Project cannot divorce itself from responsibility for possible misuse, although it might want to.

Dr. Kenneth Weiss replied that the Committee was very concerned that the data be open and not be seen as the private domain of two or three labs. Once it is open, it becomes like GenBank. GenBank is only statistical; CEPH (the Center for the Study of Polymorphisms in Paris) might be a better example because it provides physical samples and not just data. He asked whether anyone could get samples from CEPH?

Dr. Cavalli-Sforza replied that after a request, CEPH writes to members of CEPH asking for information about the scientific validity of the planned research. CEPH makes an investment in sending DNA to people for free. Its data bank is accessible to all members, but there is a screening process for getting DNA. CEPH can turn people down if their proposed work is not sufficiently interesting for the linkage project. Dr. Feldman noted that CEPH is more restrictive than the Human Genome Diversity Committee had intended the Project to be. Dr. Cavalli-Sforza was concerned about whether access should be limited or whether, if NIH funds were used, it could legally be limited. Prof. Greely pointed out that there may not be a true dichotomy between limited and open access. One could have open access for researchers to cell lines, but require information about the planned studies and advance notice of the results.

Dr. Evans noted that access involves both the cell lines and the data generated from that material, which will be in the data bank. The Project might want to have the data freely available, while demanding information about a study before giving access to samples. Someone noted that this would not require a

screening committee, but perhaps just an organization that watched how the information was being used and what issues might be arising.

Dr. Murray said the discussion of open access reminded him of the early days of sickle-cell testing, before there were any requirements for laboratory or professional certification. Many people were performing sickle-cell testing without any background, and they did lots of harm. He doesn't think the Project wants this material to be accessible to people without scientific background. If it does, one of his nightmares will come true. Someone with mischief on their mind will ask for the materials for the purpose of doing something harmful with them.

Dr. Mark Weiss said he had just been looking at the catalog of cell lines from the National Institute of General Medical Science. That catalog, he noted, qualifies the meaning of "open." Cell cultures and DNA samples are distributed only to qualified professional persons who are associated with recognized research, medical, or educational institutions, so it will not be sent to absolutely anyone. Dr. Eckstrand agreed that a researcher could not just write in from her home address, particularly because that material is limited. If the material is unlimited, such as the DNA database, then NIH has made it open to anyone. She was not sure whether cell lines should be classified as limited or unlimited. Dr. Feldman asked whether NIH imposes any restrictions. If someone writes in and says they have money from a right-wing foundation and he wants a sample, would NIH provide it? Dr. Eckstrand replied that if he writes in from a recognized institution, the bank might check with its project officer and would probably be advised to send the cell lines.

Dr. Giddings noted that the U.S. Department of Agriculture had had similar problems. He thinks the Project needs to be aware of the potential for misuse and to have contingency plans to deal with it. It should not restrict access, but it should be prepared to respond to abuses.

Prof. Greely gave as an example: the Project learns a researcher is about to claim that, on the basis of data from the Human Genome Diversity Project, all Irishmen (his own ethnic background) were inherently prone to alcoholism. In that case, the Project could have a spokesman prepared to put that claim into perspective for the press.

General Closing Discussion

Prof. Greely noted that the meeting's time was nearly up. He invited people to discuss any other issues they thought were particularly important (while adding that he would be happy to hear comments on any of those issues later).

Dr. Murray said that education about the goals and limits of the Project could help prevent a lot of problems. It was noted that there was fairly general agreement about the value of education.

Dr. Siniscalco said that classifying populations is an old habit of people. We may now be able to classify people more scientifically, but that is not necessarily a bad thing. We have moved from eugenics to euphenics, where we can try to

make the life of everyone equal regardless of what genes they are born with. To do that, we need to understand the relationship between genes and environment, which is what this Project is about.

Prof. Greely and Dr. Feldman both thanked the participants sincerely for their useful discussion of these issues, and, at 4:00, the workshop concluded.

APPENDIX A

HUMAN GENOME DIVERSITY WORKSHOP ON ETHICAL AND HUMAN RIGHTS ISSUES
FEBRUARY 17, 1993

This document contains a schedule for the workshop, a non-exclusive list of issues that may be discussed, and a listing of the workshop participants.

SCHEDULE

Within each roundtable discussion, the participants are listed in the order in which they are currently expected to speak, although those orders are subject to change. Each participant should take, at most, 10 minutes for opening comments. Most of the time during each roundtable should be used for discussion of the issues among the participants and the audience.

9:15	Opening -- Prof. Hank Greely, Stanford Law School (moderator) Summary of HGD-- Dr. Marcus Feldman, Stanford
9:30	Roundtable Discussion of Issues Raised Issues Raised by the Possible Commercial Value of the Project's Samples, Data, or Findings Dr. Val Giddings, USDA Dr. Walter Reid, World Resources Institute Dr. Jason Clay, Cultural Survival Inc.
10:50	Break
11:00	Roundtable Discussion of Issues Raised by the Sample Collection Process Dr. Ken Weiss, Pennsylvania State University Dr. Gary Ellis, NIH Office of Protection from Research Risks Dr. Rachelle Hollander, NSF Dr. Joan Porter, NIH Office of Protection from Research Risks
12:30	Lunch
1:30	Roundtable Discussion of the Possible Misuse of Project Data or Results Dr. Diane Paul, University of Massachusetts at Boston Dr. William Schneider, Indiana University-Purdue University at Indianapolis Dr. Robert Murray, Howard University Dr. Eric Juengst, NIH
3:00	Break
3:10	General Discussion of Other Issues, Including Organization Participants and audience
4:00	Scheduled Conclusion

ISSUES

The following pages contain a list of some of the issues that may be discussed at the workshop. This list, in somewhat earlier form, was circulated to the participants. The list is not intended to preclude discussion of other issues, although issues that do not relate to one of the three roundtables may be referred, at the moderator's discretion, to the afternoon's general discussion period.

I. Issues Raised by the Possible Commercial Value of the Project's Samples, Data, or Findings

This area includes a number of issues related to property, financial arrangements, and technology transfer. These issues are ethical and political. I believe we should deal with them first, because they have some implications for the second set of issues

A. Payment Arrangements in the Event the Project Has Commercial Value

1. What is the potential for the commercial use of the project's data or results?
2. Is it ethically permissible to pay, even in a contingent manner, for a human's DNA, or does that impermissibly "commodify" that person and demean his or her humanity?
3. Will some governments demand a system of payment for any commercial value from the gene samples collected by their citizens?
4. If so, how should the project respond?
5. Should the project volunteer a uniform payment system for all countries, regardless of whether they request one?
6. If the project were to agree to a kind of "royalty" arrangement, how should it be handled?
   a. Should the payment be based made on behalf of all the populations sampled or just to those whose samples contributed to the commercial value?
   b. Should the payments be made to the governments of the relevant populations, or to the populations directly?
   c. If payments were to be made to populations directly, how should that be accomplished?

B. Payment Arrangements and the Sampling Process

1. Should the payment arrangements reached, if any, or the absence of such arrangements, be part of the informed consent to individuals whose DNA is sampled?
2. Should those arrangements be subject to renegotiation by each sampled population?
3. Should the informed consent ask the sampled individuals expressly to release any property rights they may have in the samples or in the use of their DNA?
4. Is it appropriate to provide any gifts or any payment to people who agree to participate?

II. Issues Raised by the Sample Collection Process

This area includes a number of issues revolving around how the samples are taken and how information obtained from the samples is used with respect to those people sampled.

A. Who Should Be Sampled

1. Should the project ever sample children?
2. If the project will sample children, what special protections should apply?
3. How should adulthood be defined for these purposes -- by American standards, by those of the sampled population, or by both?
4. Are there other classes of potential donors who deserve special consideration?

B. Informed Consent

1. Should the project require a uniform manner of informed consent among all sampled populations?
2. If variable methods of informed consent (written, oral, etc.) are permissible, considerations should be relevant in deciding the method to use with a particular population?
3. Should the project ever accept the informed consent of one person as binding other persons (local leader and population, husbands and wives, parents and children, etc.)?
4. Should the project ever permit someone other than the person to be sampled (a local leader, spouse, parent, etc.) to veto that person's participation?

C. Privacy

1. What level of privacy should the project seek to ensure in the sampling process, with respect to the informed consent and the actual decision to donate or not?
2. Should the samples, once taken and processed, continue to be identified with individual donors?

D. Handling the Samples

1. Should the blood samples be screened for any disease organisms?
2. If so, is the screening primarily to protect people who may come in contact with the samples in the future or to help the individuals who gave the samples?
3. If any samples are to be screened for any disease organisms, who will decide what screening will be done and what considerations should that decision-maker apply?
4. If a sample is screened and is found to be contaminated with a pathogen, what should be done with the sample?
5. If a contaminated sample is found, should the person who donated the sample be told? Under all circumstances? Under some circumstances? Under no circumstances?
6. Does the project have an obligation to provide medical assistance to individual donors who are found, through the project's work, to have medical conditions?

III. Issues Raised by the Possible Misuse of Project Data or Results

This area involves mainly the possibility that the project's existence, data, or findings will be used to claim that human populations are more different than the science supports and that these differences have social or political implications. The concern about such misuse stems from two reasonable fears. First, some people may try to use the project to support racist or nationalist claims of the "genetic superiority" of particular populations. Second, well-intentioned laymen may misconstrue some of the project's findings to similar ends.

1. Based in part on the history of cultural use of genetic information, how likely is such misuse of the project?
2. Are the concerns about such misuse sufficiently great that the project should not be undertaken?
3. If this research is not undertaken through the HGD project, will it be undertaken by others in any event?
4. Should access to samples or data from the project be widely available, or should they be limited to researchers with appropriate academic credentials?
5. Should access to samples or data from the project be limited on political grounds?
6. Does the project have an ethical obligation to try to educate the public about the meaning of its results? If so, how should that obligation be fulfilled?

IV. Other Issues

This area is a catch-all for other ideas. We will discuss these are the end of the workshop to the extent we have time (and energy) left. This list currently contains only two areas of issues, but more are expected to surface during the day's discussion.

A. Technology Transfer

1. How should the project respond to a request from a host country for technology transfer?
2. Even without such a request, should the project feel an ethical obligation to attempt to transfer relevant technologies to the countries or the populations in which it samples?
3. If the project should commit to technology transfer, how can such transfer best be accomplished?

B. Organization

1. Should the project provide for some further consideration of ethical and human rights issues that it raises?
2. How should that consideration be achieved -- through an ethics advisory committee? through more workshops? through some other method?

APPENDIX E

PARTICIPANTS

The following people are participating in the workshop discussions. The extremely short biographical material has not been cleared with all of them; I regret any errors.

Dr. Jason Clay
Cultural Survival, Inc., Boston, Massachusetts. Dr. Clay has a Ph.D. in anthropology. From 1980 to 1989, he was director of research for Cultural Survival, Inc., a nonprofit human rights organization working with indigenous peoples around the world. In 1989 he founded Cultural Survival Enterprises, which he directs. In that capacity he works to develop economic resources for indigenous peoples.

Dr. Gary Ellis
Office of Protection from Research Risks, National Institutes of Health. Dr. Ellis has a Ph.D. in Biological Sciences. He recently assumed the post of director of the NIH Office of Protection from Research Risks after several years as director of the Division of Health Promotion and Disease Prevention at the Institute of Medicine.

Dr. Val Giddings
Biotechnology, Biologics, and Environmental Protection Division, Animal and Plant Health Inspection Service, U.S. Department of Agriculture. Dr. Giddings has a Ph.D. in genetics. He is Chief of Science and Policy Coordination for the branch of the USDA that regulates the use of biotechnology products in the environment. Dr. Giddings was a member of the United States delegation that negotiated the Biodiversity Treaty, and he was particularly involved in the Treaty's technology transfer and royalty provisions.

Dr. Marcus Feldman
Stanford University, Department of Biological Sciences. Dr. Feldman has a Ph.D. in mathematical biology. A leading population geneticist, he is a member of the HGD Committee.

Prof. Hank Greely(chair)
Stanford Law School. I am a law professor specializing in health policy. I have written on aspects of the Human Genome Project and on some aspects of bioethics.

Dr. Rachelle Hollander
National Science Foundation. Dr. Hollander has a Ph.D. in philosophy. She is director of the NSF Ethics and Values Studies Program.

Dr. Eric Juengst
National Center for Human Genome Research, National Institutes of Health. Dr. Juengst has a Ph.D. in philosophy. He is the principal ethicist with the NCHGR and directs its program on the ethical, legal, and social implications of the Human Genome project.

Dr. Robert Murray
Howard University, College of Medicine, Department of Genetics and Human Genetics (Chair). Dr. Murray is an M.D. He has written on genetic diversity among peoples of African origin and has long been interested in ethical issues concerning genetics.

Dr. Diane Paul
University of Massachusetts at Boston, Department of Political Science. Dr. Paul has a Ph.D. in political science and is the director of her university's Program in Science, Technology, and Values. She has written on the historical use of human diversity in the debates over eugenics and other social issues.

Dr. Joan Porter
Office of Protection from Research Risks, National Institutes of Health. Dr. Porter has a doctorate in public administration. She has served as Special Assistant to the Director in the Office of Protection from Research Risks for the past 11 years and has written on protection of human research subjects.

Dr. Walter Reid
Vice President for Program, World Resources Institute. Dr. Reid has a Ph.D. in zoology with a specialization in population and community ecology. His work has focused on world biodiversity.

Dr. William Schneider
Indiana University-Purdue University at Indianapolis, Department of History (Chair). Dr. Schneider has a Ph.D. in history. He has written on the historical uses of ethnic diversity in blood types in debates over eugenics and other social issues.

Dr. Kenneth Weiss
Pennsylvania State University, Department of Anthropology (chair). Dr. Weiss has a Ph.D. in anthropology and works in biological anthropology, epidemiology, and genetics, with a special interest in native American populations. Dr. Weiss is a member of the Human Genome Diversity committee.

Yale University

Kenneth K. Kidd, Ph.D.
Professor of Genetics,
Psychiatry, and Biology
Department of Genetics
School of Medicine
333 Cedar Street
P.O. Box 3333
New Haven, Connecticut 06510-8005

Campus address:
1-309 Sterling Hall of Medicine
Telephone: 203 785-2654
Fax: 203 785-6568

May 5, 1993

Senator Daniel Akaka
720 Hart Senate Office Building
Washington, D.C. 20510-1103
Attn: Shane Merz

Dear Senator Akaka:

Thank you for the opportunity to submit written testimony for the hearing the Senate Committee on Governmental Affairs held on April 26 to examine the potential benefits of the Human Genome Diversity Project. I am a very strong supporter of that project; indeed, I am a member of the committee that has been working to organize and obtain funding for the project. As a committee member I have participated in preparing answers to the questions you directed to Prof. Hank T. Greely in a letter dated April 28, 1993. Today, however, I wish to make a personal statement. Rather than reiterate many of the points already made in the testimony of other supporters of the project, points with which I fully agree, I will use this opportunity to emphasize additional concerns.

As a human geneticist I have strongly supported the Genome Project since its inception but always maintained that it had too narrow a focus. Genetic variation is a fundamental aspect of the genome of any organism but the priorities of the Genome Project do not include the systematic investigation of normal human variation--no research funds are allocated to the study of human variation while millions of dollars each year go to studies of the genomes of other organisms--mice, fruit flies, and round worms, for example. Those studies of other organisms are scientifically valuable, but a balance is needed. A genome is two dimensional. One dimension is the sequence, the length; the other dimension is normal variation in that sequence, the breadth. We need to study the length and breadth of the human genome to understand it.

The collection of cell lines and supplemental DNA samples that the Human Genome Diversity (HGD) project proposes to establish is designed to complement the existing Genome Project by providing that missing second dimension. Some studies of those samples are to be part of the HGD project but the samples will also be a resource for many kinds of studies that are themselves beyond the primary focus of the project. Already the small resource of cell lines that exists in my laboratory has led to medically significant findings on

normal variation. For example, normal variation in receptor molecules of pharmacological importance has been shown. This variation does not cause a disease, but can determine individual response to medications and may make it possible to predict drug effectiveness based on a simple genetic test. Some ethnic groups, such as Japanese, have variant forms of these molecules not seen in Europeans. In total we found 9 forms in Europeans and an additional 14 forms in the other populations we sampled. This is a clear example of how basic research findings in normal genetic variation can have significant clinical applications. Scientifically valuable studies, impossible to envision today, will be developed using the knowledge gained from the Genome Project and the resources available from the Human Genome Diversity Project.

The National Institute for General Medical Sciences (NIGMS) has in place a mechanism which could be expanded to deal with the cell repository aspects of the HGD project. There exists a very small collection of cell lines on diverse human populations in the Coriell Institute for Medical Research, Camden, New Jersey, sponsored by NIGMS. This collection consists primarily of a few cell lines from a few of the populations Prof. Cavalli-Sforza and I have been able to assemble over the past several years. Dr. Mulivor, Director of the Human Genetic Mutant Cell Repository, has informed me that this small collection is already one of the ten most important in the Repository in terms of the number of cell lines requested by investigators.

A final consideration that has not yet been emphasized is that of training the next generation of scientists and physicians. Understanding variation requires knowledge, skills and perspectives that are different from those required for mapping and sequencing the genome. The necessary training and experience are best achieved through participation in research projects on normal variation. Without opportunity for the study of normal variation, the perception of young scientists will become one dimensional; variation will appear abnormal rather than normal. If we do not know the range of normal, how can we recognize the abnormal?

Thank you for allowing me this opportunity to express my concerns. If I can be of any assistance in your Committee's deliberations, please do not hesitate to call upon me.

Sincerely,

Kenneth K. Kidd, Ph.D.
Professor of Genetics,
Psychiatry, and Biology

KKK:rjm

Forum

Human Genome Diversity Initiative

JUDITH R. KIDD,[1] KENNETH K. KIDD,[2] AND KENNETH M. WEISS[3]

The international Human Genome Project (HGP) is an effort to map and then sequence *the* human genome, that is, to determine the DNA sequence of all 3 billion nucleotide base pairs in the human DNA complement. This effort involves many agencies and countries and aims to identify, locate on their respective chromosomes, and sequence all functional genes, regulatory regions, and other DNA segments. Foremost among the many anticipated benefits of the project is that the data generated will lead to a better understanding of the function and expression of human genes. Unfortunately, this huge undertaking does not include any systematic consideration of *variation* in sequence; that is, the HGP does not address that aspect of the genome that is of driving interest in anthropology: human variation. It is the variation of the DNA sequence among individuals, both within and between populations, that illuminates the unique history of each population. That historical record cannot be inferred from a standardized genome sequence or from a gene map derived from a composite individual. In this forum we review some of the uses of studies of variation and then describe the general structure of a new initiative to organize an international project to address DNA variation in the human species.

The genome is not a simple homogeneous entity; it is composed of different elements, each able to yield information on aspects of our history and evolution. Molecular genetic tools allow us to learn about the different elements and hence about our past. Although all regions of the genome are composed of nucleotides, the functional content of these sequences varies. Each component reflects particular aspects of human evolution, each has a unique evolutionary tempo, and each affects and has been affected by our evolutionary history in ways different from the others. Detailed analyses of variation in all the components of the human genome will provide essential information to advance many fields of study, including functional studies, forensics, microevolution, and medical applications.

[1] Departments of Anthropology and Genetics, Yale University, New Haven, CT 06510.
[2] Department of Genetics, Yale University School of Medicine, New Haven, CT 06510.
[3] Department of Anthropology, Pennsylvania State University, University Park, PA 16802.

Human Biology, February 1993, v. 65, no. 1, pp. 1–6.

Functional Studies. Classical issues of interest that have generally resisted definitive analysis include adaptation and the relationship between genetic variation and phenotypic variation. New methods promise to revolutionize the study of these topics. Two problems confront us when we attempt to study phenotypic adaptation and control. The first is identification of the genes responsible for the phenotypic variation in question. In some cases animal studies and biochemical studies can identify the relevant loci. Often, positional cloning based on family data and linkage studies is required to identify the specific locus responsible for adaptive variation. The second problem is that mutations continually arise in all loci and most new mutations are unique at the DNA level. This means that specific adaptations will often be local, and the evolutionary history of each area will often involve different mutations, even in the same genes. Reconstruction of human adaptive evolution will require study of genetic diversity in functional genes in different populations.

Forensic Studies. There is a great deal of interest and controversy these days over the use of DNA data for forensic purposes, another of the traditional areas of interest to anthropologists. The method of DNA profiling relies on the fact that some short nucleotide sequences may be adjacently repeated a highly variable number of times. These repetitive regions are also called variable number of tandem repeats (VNTRs) or hypervariable regions (including both mini- and microsatellite DNA). Typically, a type of repeat sequence found in any given chromosomal region will be there in all persons (sometimes in other primates); however, in such regions mutation may change the copy number so rapidly that even close relatives may not share the exact same copy numbers or DNA profile for both chromosomes at several such loci.

One of the major forensic issues being debated is the degree to which the DNA polymorphisms being used for forensic identification can be interpreted when questions of small isolated populations are relevant to the specific case. The data collected so far show that for most loci the alleles have different frequency distributions in different ethnically defined samples. Those differences, however, are often small, especially for closely related populations. Even where the differences in allele frequencies are large, the *number* of alleles is usually still large in all populations. This suggests the generalization that any multilocus DNA phenotype based on these VNTR loci is a rare one on a global basis. Exactly how rare can be determined only in the context of a specific population. Currently, good estimates are possible for populations of mixed European ancestry, but those estimates do not generalize to other specific ethnic groups without either data specifically on the ethnic group or an assumption about the similarity of the ethnic group to those already studied.

Microevolutionary Studies. To date, much discussion has centered on that aspect of genetic research that concerns the reconstruction of population relationships from a phylogenetic point of view. Most valuable for this are selectively neutral sequences, that is, sequences that are likely to evolve in a more regular way as a result of genetic drift because they are not affected by natural selection. Phylogenetic relationships among populations are certainly one of the core problems in anthropology.

The hundred or so single-copy nuclear DNA polymorphisms that have been studied on a global scale (albeit in a most preliminary way) show that almost all alleles exist in almost all populations. A corollary is that most polymorphisms examined are polymorphic throughout the species. Although balancing selection might be invoked for a few loci, to invoke it for many loci is not reasonable. Therefore we must conclude that diversification of modern *H. sapiens* is recent or that it has involved relatively large effective populations at all stages or both. Polymorphisms found in all human populations are informative of our deeper history. Even when the individual polymorphisms differ only in frequency from population to population, the combinations of polymorphisms that are molecularly close—haplotypes—can be characteristic of individual populations, thus reflecting the unique history of populations (e.g., migration, genetic drift), their relationships to one another (phylogenetic relationships), and selection at nearby loci. Thus DNA polymorphisms are useful for reconstructing both long- and medium-term human prehistory.

Medical Applications. Clinical applications and medical anthropology are intertwined with population studies. The diagnosis of genetic disease, indeed even treatment, soon will depend on the identification of specific mutations at loci whose function relates to a given disease. Currently, genetic counseling, screening, and attempts at gene therapy are heavily concentrated in European-derived populations, with some exceptions for Jewish populations (e.g., Tay-Sachs disease) and populations affected with globin gene mutations such as sickle-cell hemoglobin. But for many diseases it has already become clear that the pathogenic mutations differ among populations. The mutations that cause phenylketonuria in Asia are different from those in Europe, and the same picture is emerging for every problem that is studied in genetic detail (including the abnormal hemoglobins). We know little about such variation in pathogenic mutations. Likewise, it would be difficult to screen populations in India or Mongolia, for example, for cystic fibrosis, given that most of the currently known mutations are identified in Europeans.

This situation for rare diseases caused by single-gene defects contrasts with the picture for common normal polymorphisms. The situation for more common but also more complex genetic diseases, such as di-

abetes or neuropsychiatric disorders, is less clear. Indeed, we do not even know which loci (or how many) control such diseases, but it is possible that mutations at different loci will be important in different populations. For example, susceptibility to diabetes may have a strong genetic component in Amerindians, but to understand this we must find the locus (loci). This requires family studies and polymorphic genetic markers that can be used to find linkage between marker variants at some genome region and the occurrence of disease in families. The highly polymorphic repetitive elements may be the most powerful markers for this purpose.

The Human Genome Diversity Initiative

All the studies outlined here depend on the analysis and understanding of genetic variation within our own species. A complete single composite sequence will provide little by way of answers to any of the research questions. To address the need for information on human biodiversity, the Human Genome Organization (HUGO, an international scientific society) has established an ad hoc committee to formulate plans for and then organize a project on human genome diversity (HGD). That ad hoc committee has received funding for a series of international workshops from a consortium of agencies, including the National Institutes of Health, the National Science Foundation, and the Department of Energy (in the United States), and additional support is being sought from agencies in Europe. These workshops have begun to formulate specific plans for how the HGD Project should be organized. Initially, the problem is to rationalize collection of genetic resources so that data relevant to a wide variety of research questions can utilize common resources. By the time this forum appears in print, the written documents produced by the first of the workshops should be in circulation within the anthropology and human genetics communities for comment and criticism. Some of the issues and goals are already clear.

The HGD Project aims to collect the biological resources that would facilitate, even be sufficient for, all the types of studies noted earlier. A worldwide selection of human populations needs to be chosen in a way that adequately represents our species and its history. The intention is that knowledgeable local investigators will help to identify the populations to be sampled. Then, to ensure permanent samples that can be a resource for many studies, cell lines will be established from individuals in these populations. DNA or the growing cells themselves will then be available to the world research community at no profit or perhaps even at a subsidized cost. Investigators wishing to study questions such as those mentioned here will have access to the appropriate material in the

cell-line "bank." Cultural characterization of the populations studied is of recognized importance, but extensive data collection is beyond the scope of the Human Genome Diversity Project. The genetic research will be far more valuable if cultural anthropologists are able to mount resources and study the same populations.

Exactly how the global HGD collection should be structured is a difficult and challenging problem that has at least three major aspects. First, there are the population genetics and statistical questions of how to sample: How many people should be sampled from how many local groups or geographic areas worldwide to provide an optimum resource for use now and in the future? Because many of these populations are rapidly losing their genetic distinctiveness through exogamy, it is important that the structure of the sample be appropriate; we cannot go back in 50 years to correct errors of sampling made now. Second, there are the anthropological questions of whom to sample. Should small endangered populations (i.e., those whose genetic identity or even survival is threatened) be sampled preferentially to large major populations representing predominant cultures? Which populations will provide the most information on the greatest number of questions? Third, there are many difficult issues of how to fund, maintain, and distribute such a resource and the data pertaining to it. Also, such a project should help the sampled populations, or at the least do them no harm. Useful technology transfer should also be an aim. How to address these three aspects is a difficult question.

The existence of a properly designed HGD collection can be exciting and of widespread and long-term benefit to anthropologists and all scholars interested in human evolution, history, and prehistory. The HGD ad hoc committee has organized three workshops, one to look at each set of problems. In addition to those attending the workshops, many other interested investigators have already contributed ideas concerning each problem. These workshops are preparing working drafts of how the HGD Project will be organized. These working drafts are being circulated broadly and summaries are being published. The final documents, incorporating the responses to the workshop drafts, will provide the framework for researchers around the world to apply to funding agencies and governments for financial support to collect and maintain the samples. This framework should facilitate allocation and availability of funds for this type of work, but we anticipate that actual funding will still be primarily through the awarding of research grants to individual researchers on a competitive basis.

The HGD Project can make fundamental contributions to applied and basic knowledge about human diversity. It will provide an important—many of us would say vital—corollary to the mapping of the human genome and the development of a largely European-based sequence

of the genome. Those efforts are of basic importance to anthropologists and geneticists, but we all stand to benefit in a special way from the proposed attempt to represent the genetic diversity of our species as a whole.

Received 9 June 1992; revision received 28 June 1992.

ISBN 0-16-043334-7